Tsakhur

Wolfgang Schulze

LANGUAGES OF THE WORLD/Materials 133

1997
LINCOM EUROPA
München - Newcastle

Published by LINCOM EUROPA, München, Newcastle, 1997.

All correspondence concerning *LANGUAGES OF THE WORLD/Materials* should be addressed to:

LINCOM EUROPA, Paul-Preuss-Str. 25, D-80995 München, Germany.
LINCOM.EUROPA@t-online.de

http://home.t-online.de/home/LINCOM.EUROPA

All rights reserved, including the rights of translation into any foreign language and the publication on electronic media. No part of this book may be reproduced and distributed in any way without the permission of the publisher.

Printed in Germany

Edited by U.J. Lüders

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Schulze, Wolfgang:
Tsakhur / Wolfgang Schulze. - München [i.e.] Unterschleissheim ;
Newcastle : LINCOM Europa, 1997
(Languages of the world : Materials ; 133)
ISBN 3-89586-150-2

[Languages of the World / Materials]

Languages of the world.Materials. - München [i.e.]
Unterschleissheim ; Newcastle : LINCOM Europa
Reihe Materials zu: Languages of the world
133. Schulze, Wolfgang: Tsakhur. - 1997

Printed on chlorine-free paper

Contents

Abbreviations 3

0. Introduction 4

1. General data on Tsakhur 5
1.1 Location 5
1.2 Dialects 8
1.3 Sociolinguistic informations 9
1.4 A brief look at the genetic background 10
1.5 General typological remarks 12

2. Phonology 14
2.1 Phonemes 14
2.1.1 Consonants 14
2.1.2 Vowels 18
2.1.3 Phonotactics 20
2.1.4 Phonetic processes 21
2.1.5 Prosodic features 22
2.1.6 Morphophonemics 22

3. Word formation 23

4. Inflectional Morphology 24
4.1 Noun Classes 24
4.2 Nouns 28
4.2.1 Number 28
4.2.2 Case System 29
4.3 Adjectives 33
4.4 Pronouns 37
4.4.1 Personal pronouns 37
4.4.2 Demonstrative Pronouns 38
4.4.3 Reflexives 39
4.4.4 Interrogatives 41
4.5 Numerals 42
4.6 Adpositions 44
4.7 Adverbs 44
4.8 Verbs 45
4.8.1 Stem formation 45
4.8.2 Class markers 45
4.8.3 Preverbs 47
4.8.4 Tense and aspect 48
4.8.5 Personal inflection 53
4.8.6 Participles and converbs 55
4.8.7 Infinitive and masdar 55

5. Syntax 55
5.1 Structure of the NP 55
5.2 Structure of the clause 56
5.2.1 Word order and attention flow 56
5.2.2 Semantic roles 57
5.2.3 Agreement 60
5.3 Major sentence types 61
5.4 Pragmatic functions 63
5.5 Negation 65
5.6 Questions 65
5.7 Coordination of NPs and clauses 66
5.8 Subordination 66

6. Sample text 68

7. Bibliography 71
7.1 A brief look at the history of Tsakhur linguistics 71
7.2 References 71

Abbreviations

ABL	Ablative
ABS	Absolutive
AD	"at" (locative)
AFF	Affective
ALL	Allative
AOR	Aorist
ATTR	Attribute marker
AUX	Auxiliary
C	Consonant
CAUS	Causal converb
CL	Class
CM	Class marker
COND	Conditional
CONF	Confirmed
CV	Converb
DAT	Dative
ECL	East Caucasian Languages
ERG	Ergative
ESS	Essive
FOC	Focus
FUT	Future
GEN	Genitive
GER	Gerund
hum	human
I,II,III,IV	Noun Class (inherent: (I), (II), (II), (IV))
IMP	Imperative
IMPERF	Imperfect(ive)
IN	"in" (locative)
INF	Infinitive
LOC	Locative (undifferentiated)
m(ask)	masculine
MASD	Masdar (verbal noun)
NARR	Narrative
NEG	Negation
NP	Noun phrase
OBL	Obliquus
PART	Participle
PAST	Past tense
PL	Plural / Proto-Lezgian
PP	Postposition
PRES	Present tense
PROH	Prohibitive
PV	Preverb
Q	Question particle
R	Sonant
REFL	Reflexive
SA	Stem augment
SAP	Speech Act Participant
¬SAP	Non SAP
SEC	South East Caucasian
SECL	South East Caucasian languages
SG	Singular
SUB	"under" (locative)
SUPER	"over" (locative)
TAM	Tense, aspect, mood
TOP	Topic
TV	Thematic vowel
V	Vowel

0. Introduction

Tsakhur is spoken by some 13.000 people who dwell in about 30 villages or settlements at the headwaters of the river Samur (valley of *Gorgin Magal*) in Southern Daghestan. This area is confined in the west by the Guton and Gjultydag mountains, in the north by the Alaxundag mountain range (the southern part of the Samur Range). Towards the east the area touches a Rutul speaking population near Luček. An important group of Tsakhur speakers can also be found on the southern slopes of the Kjabak-Tepe range in Northern Azerbajdzhan (along the two tributaries of the Agri-Čaj river (Katex-Čaj and Kurmux-Čaj) and south of them). "Tsakhur" is the (somewhat disputed) name for a dialect continuum, that is named for the village of Tsakhur (in the Samur valley). Together with Rutul, the language forms the western branch of the Samur languages, itself being a subgroup of **South East Caucasian (Lezgian)**. Though Tsakhur is only sporadically written (a new "written language" has recently been reintroduced), it is quite vivid in ordinary life (competing especially with Azeri).

Tsakhur is a "typical" Lezgian language, operating through a system of (partly) semantic ergativity and noun classification, based on extensive case marking and a complex verbal paradigms. Though Tsakhur is heavily agglutinating, inflectional features can often be observed. As opposed to some other Lezgian languages, Tsakhur shows a tendency towards personal agreement (restricted, however, to the first person), ergative case marking of personal pronouns (in Gelmets), and the development of focus particles.

The booklet informs on the basic structure of Tsakhur (phonology, morphology, morphosyntax, morphosemantics, and syntax), which is (at least partly) explained on the basis of internal and external reconstruction. But it should be born in mind that the scope of this book does not allow a comprehensive treatment of all features of Tsakhur grammar. This limitation is especially valid for the sections on verb morphology and on syntax. Many aspect of Tsakhur syntax such as word order with relation to attention flow and view point, pragmatic functions (esp. focusing strategies), and personal agreement patterns have had to be treated quite rudimentarily. Much more work has to be done on these problems before a more complete picture can be drawn. Still, the informations given cover most of the prototype of Tsakhur syntax. Finally, a sample text (taken from Ibragimov 1968) together with interlinear glosses helps to illustrate the linguistic structure of the language.

The dialect covered by the present booklet is that of the village of Caxur (*ts'aˤχ*) which is the perhaps most representative variety of the Tsakhur dialect continuum. However, if necessary, reference is also made to the dialects of Sabunči (Sabunchi) and Gel'mec (Gelmets).

The language data stem either from written sources mentioned in the bibliography or from my own fieldnotes gathered during three stays in Makhatchkala and Southern Daghestan. The written sources have been checked with informants during these stays (except for Ibragimov 1990, but most of the data given there conform to my own field notes). The Tsakhur phrases quoted in the text are taken from Ibragimov 1968, ʒeiranišvili 1983/84, Ibragimov 1990 as well as from my own field notes. In order to keep the text within the limts of the LW/M series no specific references for the single citations have been given (also because many of them can be easily checked with the help of Ibragimov's publications).

As for transcription and transliteration: Place names etc. are given as transliterations of the corresponding Russian names, language names, however, appear in their English form. Language data are given in a broad IPA version. Due to the fact the pronunciation of Tsakhur

is conditioned by a remarkable degree of idiosyncrasy, lexical entries may vary within the text of this booklet. These varieties are especially related to the degree of palatalization and pharyngealization of lexical forms. Consequently, every lexical entry is quoted just as it has been noted in the field or found in the descriptions of Tsakhur. No efforts have been made to normalize them.

Linguistic research on Tsakhur started with some (lexical) entries in Erckert 1895 which had been exploited by G. Hüsing in order to substantiate his claim that Elamic originated from Tsakhur (Hüsing 1910). In 1913 A. Dirr published the first grammar on the language (Dirr 1913), the material of which was used by N. Trubetzkoy is his famous article on the consonant systems of East Caucasian languages (Trubetzkoy 1931). The short period of written Tsakhur (1934-38) saw some (pedagogical) works by A. N. Genko and S.A. Džafarov (cf. 1.3). Later, Ev. ʒeiranišvili treated the language in the light of his theory that Tsakhur and Rutul form a close dialectal continuum (culminating in ʒeiranišvili 1983/84), whereas G. X. Ibragimov (a native speaker) concentrated on the language itself which he treated for its phonetics in Ibragimov 1968. Ibragimov 1990 perhaps is the most comprehensive (and best) treatment of Tsakhur at all. The present booklet represents the first consideration of the language from outside the former Soviet Union (as far as I know). The reader should bare in mind that it is written by a non native (English) speaker, so some awkward expressions etc. may occur. I hope that this fact does not effect the description of Tsakhur grammar itself, which is worthwhile to be studied in a much more comprehensive way than it could have been done here.

1. General data on Tsakhur

1.1 Location

Tsakhur has some 13.000 speakers who dwell in about 30 villages or settlements situated at the headwaters of the river Samur (in the valley of *gorgin magal*) in Southern Daghestan (Russia). An important group of Tsakhur speaking people can also be found in Nothern Azerbajdzhan (between the two tributaries of the Agri-Čaj river, namely Katex-Čaj and Kurmux-Čaj). They represent one third of the Tsakhur speaking population. The following villages and auls are inhabited by Tsakhur speakers[1]:

Official name	Tsakhur name	Mainly Tsakhur Speakers	Mixed with
Daghestan:			
Kurdul	(ɢurdul) lek	yes	
Gel'mec	gelmets', gemits', gimits', gelmez,	yes	

[1] Tsakhur speakers in major towns such as Baku, Maxačkala etc. are not mentioned in this list, which gives only the names of those villages and auls that constitute the geographical area of Tsakhur. It should be born in mind, however, that in the larger villages Zakatal' (*zagatala*) and Kaxi (*gah*) (Northwest Azerbajdzhan) Tsakhur speakers form an important minority.

	k'irmits', gəlʲmets', gəblits'		
Mikik	mix:eʏ, mixeʏ	yes	
Xijax	χəˢje, χəˢjaˢʁ	yes	
Caxur	ts'aˢχ	yes	
Sjugjut	sogʲut, søgyt, χojik	yes	
Muslax	məslaχ, muslaʁ, məslaʁ, mətslaq	yes	
Mišleš	məʃleʃ, miʃleʃ	yes	
Korš	k(')orʃ	yes	
Džynyx	dʒiniχ, dʒinəʁ, dʒinaˢʁ, dʒinij, dʒini(j)	yes	
Attal	at:'al	yes	
Kaljal	kalelʲ	yes	
Baš-Kaljal	baʃkalelʲ	yes	
Baš-Muxax	(baš-)məˢχaˢχ	yes	

Azerbajdžan:

Agdam-Kaljal	aʁdam-kalelʲ	yes	
Suwagil'	səwagəlʲ, suwagyl suwageł, suwagøl, səwagəl	yes	
Nižne-Suwagil'	k'as, q'as	?	
Jeni-Suwagil'	teze-səwagəlʲ		Azeri
Karkay	k'ark'ʷaj, q'arq'ʷaj, q'arq'aj	yes	
Kalalu	ɢalal, q'um	yes	
Sabunči	sap'untʃi, sap'əndʒi	yes	
Alaskar	ælæsk'ær, halasq'ar	yes	
Muxax	məˢχaˢχ		Azeri
Mamrux	mamroj, mamrij, mamərі(j) / mamrəχ	yes	
Gezbarax	gølbaraχ, gøzbaraχ, gudbaraχ, gutʃbaraχ, gudʲbarag	yes	
Mišleš (Činčar)	tʃ'inar	yes	
Ljakit-Ketjuklju	lekit, lækit(-køtykly)		Azeri
Emirdžan	---		Azeri
Uzjumlju[2]	uzymly		Azeri
Alibajramly	---		Azeri, Udi

[2] Uzymly is a quite recent settlement founded by some forty families from Caxur in the late sixties.

Azgil	æzgil	?	
Čobankol	---		Azeri
Alibeglo	alibejlo		Azeri
Džimdžimax	---		Azeri
Čardaxlar	tʃaˢrdaχbə		Awar
Tala	tala		Awar, Azeri
Gjulljuk	gullʲuk		Azeri
Zarna	zarna		Azeri

The actual distribution of these villages and auls can be seen as the reflex of older economical traditions: The Tsakhur people mainly were sheep-breeders, using the pasture-grounds of the Samur valley in summer times only. In October they left these pastures and, using the passes of Dindi, Attagaj, and Baljakada, they moved south up to the river Iori and to the left bank of the river Kura. Consequently, they occupied the vast region of the Alazani valley, the Iori plateau, and the Džejrančel' plains. Historically, Tsakhur settlements in Azerbajdžan could be found down to the region of Mingečaur in the East and Dedoplis-Ckaro in the West. Later, however, the inhabitants of the more southern settlements have switched to Azeri, whereas those of the villages in the Alazani valley maintained their native language. Thus we can hardly tell what the original "homeland" of the Tsakhurs has been: We may think of the Alazani valley and the Iori plateau to be the fist candidates, also because these regions show a much more developed economy (horse breeding, gardening etc.). However, the linguistic links of Tsakhur especially with Rutul (spoken nearly exclusively in the Samur valley (cf. fn.7)) hint at an intermediate "homeland" north of the Great Caucasus mountain range. We have to assume that the speakers of Proto-Rutul-Tsakhur once dwelled just in those places in Nothern Azerbajdžan that are now (again?) occupied by Tsakhur speakers. Here they constituted one part of the Proto-Lezgian (later Proto-Samur) dialect continuum. The region of the Alazani valley and its surroundings are now thought to represent even the "Urheimat" of Proto-East Caucasian[3].

The strong ties between the Samur and the Alazani regions is also documented by the fact that historically the villages mentioned above formed six "free communities", five of which disregarded the mountain barriers:

1. The community of Gel'mec (Gel'mec and Lek)
2. The community of Caxur (Caxur, Xijax, Mikik, Sjugjut, Muslax; Kum, Činar, Ljakit- Ketjuklu, Zarna...)
3. The community of Mišleš-Tala (Mišeleš; Tala, Činčar...)
4. The community of Džynyx (Džynyx, Korš; Gjulljuk, Gezbarax)
5. The community of Muxax (Baš-Muxax, Kjalal; Agdam, Kjalal, Sabunči, Muxax...)
6. The community of Suwagil' (Suwagil', Kas, Karkaj...)

[3] Cf. Schulze 1997a and especially Nichols 1997.

These (partly feudal) communities which were also characterized by a different degree of endogamy can be seen as forming the starting point for the dialectal differentiation of Tsakhur (cf. below). Due to collectivization and the administrative separation of the two territories (Daghestan vs. Azerbajžan) these free communities have become distorted, contacts between the villages have slowed down, and - as a result - regional dialects developed.

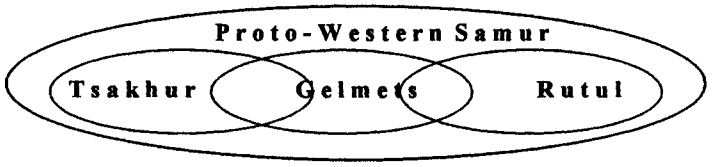
1.2 Dialects

"Tsakhur" is the (somewhat disputed[4]) name for a dialect continuum, that is named for the village of Caxur (in the Samur valley), a common self-denomination does not exist[5]. Today, the dialect continuum has two extremes, namely Tsakhur proper and Gelmets. Whereas Gelmets is exclusively spoken in Daghestan (Gel'mec and Lek (Kurdul)), Tsakhur has many dependencies in Azerbajdžan, which can be grouped according to subdialects:

1. Tsakhur-Kum[6]: Caxur, Xijax, Sjugjut; Kum, Činar, Ljakit, Uzjumlju, Kaxi;
2. Tsakhur-Mišleš: Mišleš; Tala, Zakatal';
3. Džynyx: Džynyx, Korš, Attal; Alaskar, Mamrux, Gezbarax, Gjulljuk;
4. Muxax-Sapunči: Muxax, Kaljal; Sapunči;
5. Suwagil': Suwagil', Karkaj, Kas, Kalal; Zakatal' (region).

The differences between the Tsakhur subdialects are marginal, though they may identify the provenience of a speaker. However, Tsakhur and Gelmets differ considerably: Gelmets is sometimes thought to have a Tsakhur grammar/lexicon with Rutul phonetics. In fact, Gelmets is the relict of a transitory dialect within the old Tsakhur-Rutul continuum ("Proto-Western Samur", cf. below):

(1)



The strong ties between Gelmets and Rutul[7] resulted in the claim, that both Tsakhur and Rutul form a single language (thus e.g. ǯeiranišvili 1983). However, according to my informants,

[4] Cf. the discussion by ǯeiranišvili 1983:23-29 and Ibragimov 1990:13-15.

[5] The name *ʦ'aˈχ* can be analyzed as on old locative (*-χ) of **ʦʷ'a* which means "fire" or "hearth". It also denotes "group of people using a common hearth". Place names in -χ are typical for Tsakhur (cf. *məslaχ, ʤiniχ, məˈχaˈx, mamrəχ,* and *gəlbaraχ*.

[6] The dialectal area of Tsakhur-Kum corresponds to the extension of the former sultanate of Cax (or Ilisu).

[7] "Rutul" is the name for a group of dialects the 12.000 speakers of which dwell along the rivers Samur, Korš-Samur, Kara-Samur, and the headwaters of Axty-Čaj and Šin rivers in the southeast of the Republic of Daghestan. Furthermore, four villages are found in Northern Azerbajdzhan. The denomination *Rutul* is based

even between Gelmets and Rutul a mutual intelligence is hardly given. Many people in Gel'mec and Lek (Kurdul) claim that they speak both Rutul and Gelmets. In fact, the differences between Gel'mets and Rutul are as great as those between Gelmets and Tsakhur proper (though mutual intelligence between Tsakhur und Gelmets speakers still is given).

Tsakhur proper differs from Gelmets also with respect to the degree of foreign influence. Whereas Gelmets is influenced only from Rutul (and partly from Azeri), the Tsakhur subdialects show language contact with Azeri, the Zakatal' and Kusur dialects of Awar, and the Ingiloi dialect of Georgian. The Sabunchi subdialect of Tsakhur has perhaps undergone the strongest pressure from Azeri, which accounts for the fact that Sabunchi is sometimes regarded as a separate dialect.

1.3 Sociolinguistic informations

Like many other indigenous East Caucasian languages, Tsakhur hardly has any written tradition. There are hints at a possible use of Tsakhur to translate some medieval Arabic sources (in the 13th century[8]). Unfortunately, no such documents have ever been found. However, there are (esp. in the villages of Caxur and Kurdul) some inscriptions on grave stones (dating from the 15th century and later) that contain at least some authentic Tsakhur names as well as some (very short) informations on the person buried, which can be regarded as Tsakhur (written in Arabic). Lateron, either Arabic, Osmanic, or Russian have been used as a means for written communication (as well as in education).

In 1931 efforts have been made to introduce a literary language based on the dialect of Tsakhur[9]. Since 1934 Tsakhur had been taught in primary schools, but these efforts were abandoned in 1938. In 1989 the language once again obtained the nominal status of a literary language though its use seems to be very limited at least until now.

In everyday life Tsakhur speakers use Russian in Daghestan and Azeri in Azerbajdžan as their written language (which corresponds to the language of education in each area). Still, Tsakhur is quite vital, not only as a language of communication within the family, but also as a means to communicate among the villagers. The Tsakhur lexicon covers the needs of both the material culture of traditional Tsakhur society and the modern economy (using many loans from Russian and Azeri). In 1979, 95.1% of the Tsakhurs claimed Tsakhur to be their first

on the name of the economically most important village of Rutul at the Samur river. Together with the speakers of Muxad, Kufa, Kiče, and Xnjux (in the east of the area) it constitutes the largest dialectal variety, called Muxad (5.000 speakers). Further north the dialect of Shinaz is represented by the settlements of Šinaz and Una (some 2.000 people). A third dialect (west of Shinaz) is spoken in the villages of Murxek, Cudikde (also Novyj Muxrek), Nacma, Fartma, and Džilixur (2.600 speakers). A quite different dialect is spoken in Ixrek in the northwest. Another important dialect is that of Borch-Khnov, situated in the mountains of Kjabjak-Tepe (in the villages of Borč and Xin or Xnov). Cf. Ibragimov 1978 and Maxmudova/Van den Berg (in preparation) for details.

[8] E.g. Zakariyā al-Qazwīnī, see Šixsaidov 1987.

[9] See Genko 1934. Some important documents are the publications of S. Çəfərov ("Alifbej" (Baku 1935), "Ʒettijn kitab" (Baku 1936), "Qədesin kitab" (Baku 1936), "ədəbijjatьn xrestomatija" (Baku 1937), and the school grammar "Ʒəxni mizen grammatika va oḳanasьn qajdabь" (Baku 1937)).

language (compared to 96.6% in 1970). Contrary to the speakers of some other Lezgian languages (like Rutul or Aghul), statistics show an increasing use of Tsakhur even among urbanized people (cf. 88.2% in 1970 against 57.7% in 1959). The recent introduction of radio broadcasting in Tsakhur (though some hours the week only) together with efforts to publish Tsakhur periodicals reflect this increasing interest in the Tsakhur language.

As it can be observed in many (especially southern) East Caucasian languages there is a strong tendency among Tsakhur youngsters to use a modified Tsakhur vernacular mixed with both Russian and Azeri. This language is characterized by a reduction of the grammatical inventory (class markers, tense forms etc.), by an "azerification" of Tsakhur phonetics (e.g. by substituting pharyngealized vowels through palatals), and by introducing Russian and Azeri phraseologies. However, the reputation of the parents' language may prevent this "young people's Tsakhur" to become the future means of overall communication.

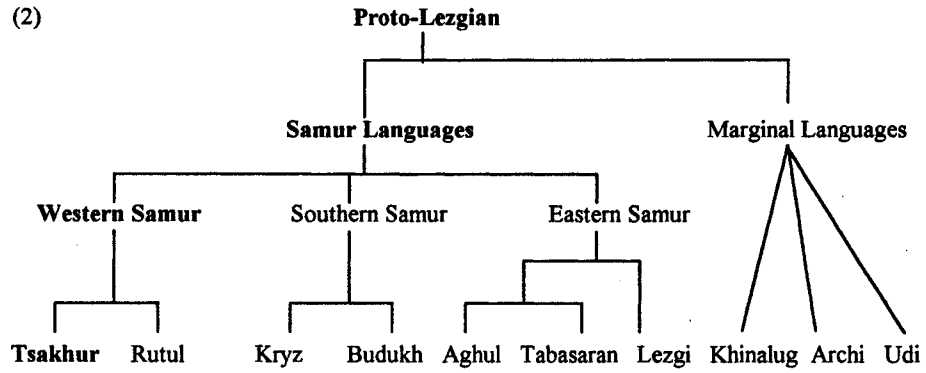
1.4 A brief look at the genetic background

Tsakhur belongs to the southern (Lezgian) branch of the East Caucasian language group. The nine or ten Lezgian languages which form this southern branch (also called "South East Caucasian languages" (SECL)) are spoken in a quite compact area that is centered by the river Samur. Reflecting their closer genetic relationship we can subdivide this group into three smaller subgroups (forming the "Samur" branch of the SECL[10]), joint by three marginal languages the speakers of which seem to have left the area of the common proto-language quite early. Most of the languages are quite vital and in everyday use. The Samur branch consists of seven languages, which are clearly interrelated. Three subgroups can be established: Eastern Samur (three languages), Western Samur (two languages), and Southern Samur (two languages). The following schematic "stammbaum" may clarify the position of Tsakhur within the SECL group[11] (it should be noted that this „stammbaum" does not reflect migration processes within the Samur branch. Moreover, the subsumption of Archi, Xinalug, and Udi under the label „Marginal languages" is not intended to signalize any closer relationship among them. The term is only used to separate these (abherent) languages from the other members of the language family):

[10] The term "Samur" is taken from Dirr 1928, though he includes Udi and Khinalug which is (at least geographically speaking) incorrect. Naturally, "Samur" reflects the actual geographical location of the languages in question only (which is problematic regarding Kryts and Budukh). Historically seen, all Lezgian languages now spoken along the Samur river result from the migration of the speakers of Proto-Samur (or Proto-Lezgian) into this area from the South of the Great Caucasian mountain range (cf. fn.3).

[11] There are some important hints at a possible closer link of Western and Southern Samur, though the details are still not very well understood. Interestingly enough Tsakhur (partly together with Rutul) shared some remarkable isoglosses with Udi. We may assume that these isoglosses are due to language contact during a time when Proto-Tsakhur-Rutul still was spoken in a region in neighborhood with the Udi area in Nothern Azerbajdžan (that is, northwest of the Agri-Čaj river).

(2)



The question whether all southern Daghestanian languages belong to a single language group genetically, which then could be interpreted as one of the offsprings of the East Caucasian proto-language (PEC), has been debated since long[12]. Today, the subgrouping of the Samur languages as shown above can be substantiated by well-established sound laws both for lexical items and morphemes. They hint at the fact that the Eastern and Southern Samur languages have undergone important innovations, whereas the western branch seems to be more archaic. The basic characteristics of Proto-Samur are best conserved in Western Samur, and we may conclude, that Proto-Samur was very much like Proto-Tsakhur-Rutul. If we include the three marginal languages, however, the picture changes dramatically (both in phonetics and morphology). The status of Archi as "Lezgian" seems now to be no longer disputed[13], though its obvious Awaro-Andian "look" has not yet been sufficiently explained. As had already been said, Archi is still more archaic than the Western Samur languages, and when we include it into our considerations regarding the type of Proto-Lezgian, we have to conclude, that all Samur languages share common innovations which have changed this type in many respects[14]. Udi, on the other hand, has not been as strongly disputed as Khinalug or Archi. The more people knew about Udi, the more came clear that it is a Lezgian language which left the dialect continuum of the proto-language quite early (but cf. fn.11). In its isolation, Udi has perhaps been the most innovating of all Lezgian languages. Khinalug, however, seems to be so far away from the Lezgian prototype that many researchers had and still have serious doubts concerning its Lezgian character. Yet, in some respects Khinalug can serve as a "missing link" in order to substantiate some claims on Proto-Lezgian that else can only be made hypotheti-

[12] For a more detailed account see (among others) Talibov 1980, Alekseev 1985, Schulze 1988a, Schulze 1997a, and Nikolayev/Starostin 1994 (with some problematic claims).

[13] Earlier, Archi was sometimes seen as an Awaro-Andian language (thus Trombetti 1922:128, Mikailov 1968, 1972, somewhat skeptical Deeters 1963:11) or as a "mixture" of Awar, Lak, and Lezgian languages (Dirr 1928, Bokarev 1961:32, Giginejšvili 1977). Trubetzkoy 1922:186 gave Archi the status of an independent East Caucasian ("tchétchène-lesghienne") branch.

[14] With respect to phonetics, Proto-Samur is characterized by the total loss of laterals (see below), which Archi has retained because of its strong convergency towards Awar. In morphology, the Samur languages have leveled the semantic distinctiveness of noun classes as well as the techniques to mark aspectual categories.

cally[15]. When we consider the obviously late arrival of the Khinalug speakers in their present settlement the abherent character of Khinalug described e.g. by Alekseev 1984 und Alekseev 1985[16] can be explained in two ways: Either: Khinalug was a Lezgian language which met a heavy substratum in the new surroundings of the Shah-Dagh mountains, which itself is not yet identified. Or: Khinalug was not Lezgian in ancient times but was "lezgified" lateron in the Shah-Dagh. Both proposals have their pros and cons. However, there are some features in Khinalug which at least remind us of the Nakh languages (Chechen, Ingush, and Bats), and which cannot be explained through later convergencies. Thus, the status of Khinalug could best be described as being "lezgoidic", though it remains unclear, whether the "Lezgian" characteristics of the language are due to divergency or convergency.

1.5 General typological remarks

The Lezgian languages share many of the typological features that are claimed to be characteristic for all East Caucasian languages. In phonetics we can observe the typical series of glottalized and/or rounded stops and affricates, an extensive use of velar, uvular, and pharyngeal features, in morphosyntactic terms we can speak of the Lezgian languages as being heavily agglutinating, operating through complex case systems with extensive distinctions within locatives and through a quite differentiated system of verbal inflection. The "operating system", that is the canonical way of construing simple sentences, is ergative, though not as much semantically or pragmatically motivated as that of other typical ergative languages. In fact, Lezgian languages are very much accusative with respect to syntax and pragmatics. The principle of noun classification which was dominating nearly all morphosyntactic categories in the proto-language has been formalized or reduced to a mere formal criterion in many languages. Lexically, all languages have suffered from heavy pressure by neighboring languages, starting in ancient times with Classical Armenian (especially important for the southern languages) and Iranian (Middle Persian). Later, Arabic loans entered the lexicon via Persian and Turkish which themselves also shaped the present lexical inventories. Nowadays it is Azeri which not only serves as an important source for new words but also exerts a strong influence on the morphosyntax and the phonological systems of the languages in question.

Though Tsakhur can be regarded as being a representative of the "Lezgian" type of languages, it yet shows some grammatical particularities, which sets the language somewhat

[15] The status of Khinalug can thus be compared to the role that Hithite plays in the reconstruction of Indo-European, cf. Schulze 1988a, Schulze-Fürhoff 1994a. The first comprehensive treatment of Khinalug as Lezgian is given by Talibov 1960. Nikolayev/Starostin 1994 opt against the inclusion of Khinalug into the Lezgian group.

[16] In a lexico-statistical analysis Alekseev 1985:23 tries to show that Khinalug shares only 22% (with Lezgi) up to 28% (with Tabasaran, Budukh, and Udi) of its basic vocabulary with the other Lezgian languages (as opposed to rates between 40% and 69% (Tabasaran-Aghul) elsewhere in this group). However, Alekseev disregards many important sound laws which would increase the rates considerably, cf. Schulze 1988a. - For Tsakhur, Alekseev gives the following figures: Lezgi 38%, Tabasaran 41%, Aghul 45%, Rutul 50%, Kryts 38%, Budukh 36%, Archi 34%, Udi 33%, Khinalug 24%. The rate for Rutul is quite remarkable, because it seems quite low compared to the strong correspondencies in grammar. We may conclude from that that Tsakhur has met a substrate (or adstrate) different from that of Rutul (basically some representatives of Kartvelian and/or Armenian).

apart from its sister languages. On the one hand ablaut phenomena due to specific morpho-phonological processes are much more salient in Tsakhur than in any other Lezgian language (except for Rutul). Ablaut effects noun inflection as well as verbal paradigms. On the other hand Tsakhur has developed some specific grammatical categories, unknown to most of the other languages: First, Tsakhur knows a complex system of indexing attributes (and locatives) according to pragmatic features (Determination, focus etc.). Second, the language has introduced a technique of marking the first person ("EGO") on the verb, which is unique within the language group. It is based on the attributive morphology and restricted to the (topicalized) first person. The other Lezgian languages either follow the canonical type of verbal agreement (via class markers), or they do not know any agreement technique at all (Aghul, Lezgi). Only Tabasaran and Udi have a system personal inflection, which in Tabasaran is restricted to Speech Act Participants. Contrary to Tsakhur, both paradigms are based on the agglutination of enclitic pronouns which at least partly serve to encode focus (or topic) functions.

The morphosyntax of Tsakhur has a considerable semantic component which is especially relevant in the case system: The categorial procedure of noun classification is intertwined with those grammatical functions which a noun covers in a sentence. The most important classification is that of "control": This prototypical feature can have its specific radial extensions depending on which role a classified noun plays. Moreover, the ergative strategy to encode simple sentences is highly role dominated, referring to semantic roles rather than to aspects of subject assignment. Consequently Tsakhur lacks (as far as we know) any strategy of perspectivization (passives or antipassives). Even the so-called "binominative" construction[17] common to some (S)ECL is completely unknown. However, the semanticity of case marking and agreement tends to decrease and to become a mere grammatical feature even in Tsakhur. Instead, pragmatic features such as topicalization or focusing gain some considerable importance.

[17] A term which is quite misleading in the context of ergativity. An alternative could perhaps be "foreground focusing" or "doubled subject construction", because the strategy of "binominatives" involves the foregrounding of the transitive agent without backgrounding the patient, e.g. Lak *pːu bawχunu ur ʧu* (father.ABS III.buy.AOR I.AUX.PRES.3.SG horse.ABS) "Father has bought a horse" (instead of *bːutːal bawχunu bur ʧu* (father.ERG III.buy.AOR III.AUX.PRES.3.SG horse.ABS)).

2. Phonology[18]

2.1 Phonemes

2.1.1 Consonants

Tsakhur (dialect of Tsakhur) shows the following system of consonants[19]:

(3)	Stops				Affricates				Fricatives		
	vl			vd	vl			vd	vl		vd
	pl	ej	gem		pl	ej	gem		pl	gem	
1	p	p'	pː	b							w
2									f		
3	t	t'	tː	d	ts	ts'	tsː	dz	s	sː	z
3p	[tʲ]	[tʲ']		dʲ	[tsʲ]	[tsʲ']			[sʲ]	[sːʲ]	[zʲ]
3l	tʷ	[tʷ']			[tsʷ]	[tsʷ']					[zʷ]
4					tʃ	tʃ'	[tʃː']	dʒ	ʃ	ʃː	
4l					tʃʷ	tʃʷ'	[tʃːʷ]	[dzʷ]	[ʃʷ]	[ʃːʷ]	
5	k	k'	kː	g					x	xː	ɣ
5p	kʲ	[kʲ']	[kːʲ]	gʲ					xʲ	[xːʲ]	
5l	[kʷ]	[kʷ']		[gʷ]					xʷ	[xːʷ]	
6	q	q'	qː	ɢ					χ	χː	ʁ
6l	qʷ	qʷ'	[qːʷ]	[ɢʷ]					χʷ	[χːʷ]	
	[ʁʷ]										
6ph	qˤ	qˤ'	[qːˤ]	ɢˤ					χˤ	[χːˤ]	ʁˤ
7		ʔ							h		
7l									[hʷ]		
7ph		ʕ'							hˤ		

Nasals, liquids, and glides: m, n, [nʲ], l, lʲ, r, y.

The phonemes in question partly show a restricted distribution. Geminates only occur in intervocalic position. *ɣ* is quite rare, *ʁ* is rare in the beginning of a word. *ɢ* appears only initially and in intervocalic position, being replaced by *q* in final position. Labialized phonemes do not appear as systematic as e.g. in Tabasaran or Aghul, but some of them have clear phonemic status. The strong tendency to use palatalized sounds is typical for Tsakhur, palatalization is not bound by equivalent contexts. Though minimal pairs are sometimes difficult to establish,

[18] A quite comprehensive (though partly controverse) treatment of Tsakhur phonology is given by Ibragimov 1968, cf. also ǯeiranišvili 1983, Ibragimov 1990:18-54. The phonemic systems of all SECL are monitored briefly by Kibrik/Kodzasov 1990:335-347. Diachronic aspects with respect to Proto-Lezgian are discussed by Talibov 1980, Schulze 1988a, and Nikolayev/Starostin 1994:139-140.

[19] 1 labial, 2 labiodental, 3 dental, 4 palatals, 5 velar, 6 uvular, 7 laryngeal; l labialized, p palatalized, ph pharyngealized; vl = voiceless, vd = voice; pl = plain, gem = geminate, ej = ejective.

their phonemic value seems to be transparent (those sounds the phonemic status of which is not yet proved are shown in square brackets).

What we in fact can describe is the prototypical system of Tsakhur consonantal phonemes. It may have its idiosyncratic or regional extensions as well as modifications, the functional status of which still is not fully clarified. Below examples for those consonants are given that belong to the prototypical system[20]:

b (i,m,f): *balag* "bag", *bəzbəzaj* "fly", *abaj* "mother-in-law", *eb* "blood";
p (i,m,f): *pifni* "lip", *pəχʲ* "saliva", *ɢaraʁ* "small roof", *kup* "container";
p' (i,m,f): *pilp'ili* "little lamp", *sap'ən* "soap", *sip'* "nip";
pː (m): *t'opːi* "small container", *ts'epːa* "first", *χepːa* "much";
w (i,m,f): *waz* "moon", *werəʁ* "sun", *q'awa* "table", *χiw* "village"[21];
f (i,m,f): only in loans[22]: *fəkər* "thought", *təfang* "gun";
t (i,m,f): *tesːi* "wooden dish", *tawasar* "pan", *top* "ball", *aχtəna* "high", *lit* "burka";
t' (i,m,f): *t'et'* "flower", *t'əb* "finger", *sat'əl* "steal", *q'at'* "beam";
tː (m): *abatːej* "father-in-law", *atːa* "(thrashing) floor", *xatːe* "wooden trough";
d (i,m,f): *dama* "river", *diχ* "son", *adi* "grandmother", *ad* "inside", *dadal* "cock";
ts (i,m,f): *tsəbelʲi* "autumn", *tsəts'ʷa* "grave stone", *jats* "bull", *məts* "wind";
ts' (i,m,f): *ts'eʔ* "goat", *ts'a* "fire", *muts'ul* "thread", *nats'* "reed";
tsː (m, rare): *kʲotsːes* "to cast", *otsːe* "hive";
dz (rare): *gandz* "rock", *dzidziler* "sort of khinkali", *odza* "stick";
s (i,m,f): *sa* "one", *sili* "tooth", *səwa* "mountain", *sos* "bride", *q'asda* (def.) "old";
sː (m): *məsːə-* "hungry", *tesːi* "wooden dish", *adesːe* "scissor";
z (i,m,f): *zer* "cow", *gəzgalʲi* "itching", *waz* "moon", *jiz* "snow";
dʲ (i,m): *dʲidʲej* "grandmother", *adʲi* "grandfather", *gadʲe* "boy", *dʲolez* "near";
tʷ (m,f): *itʷ* "honey", *getʷi* "hit";
tʃ (i,m,f): *tʃodʒ* "brother", *χotʃe* "snake", *itʃi* "girl", *etʃ* "apple", *gatʃ* "horn";
tʃ' (i,m,f): *tʃ'ar* "hair", *tʃ'ulʲkʲuˤ* "clay", *itʃ'es* "to enter", *k'əʧ'* "button";
tʃː (m,rare): *q'atʃːi* "hat", *satʃːes* "to close";
dʒ (i,m,f,): *dʒilʲi* "floor", *dʒod* "rough", *dʒedʒim* "tapestry", *uˤldʒuˤm* "week", *sadʒ* "frying-pan", *gardʒ* "splinter", *oˤdʒ* "wickedness";
ʃ (i,m,f): *ʃi* "we", *ʃu* "you", *jiʃda* "our", *ʃiʃ* "spit", *q'əʃ* "butter";
ʃː (m): *xunaʃːe* "woman", *waʃːaq'* "leopard", *boʃːu-* "deaf";
tʃʷ (m,rare): *hetʃʷi* "sowing", *gʲetʃʷi* "cooking of meat";
tʃʷ' (rare): *itʃʷ'i* "entrance", *g'etʃʷ'i* "chewing";
k (i,m,f): *kəts'a* "leg", *kun* "dish", *kok-* "fat", *kuklak* "castle", *dakruk* "roller";
k' (i,m,f): *k'at'e* "chicken", *k'uk'* "spoon", *k'osːa* "cough", *k'əlʲk'am* "liver";

[20] The abbreviations "i,m,f" denote the possible position of each phoneme within a word (initial, medial, final). However, the distribution indicated below is not always balanced statistically.

[21] When followed by a patalal vowels, *w* tends towards a labiodental realization ([v]): Thus we have *χiwe* ~ *χive* "in the village".

[22] *f* is always secondary in the Lezgian languages. In Tsakhur, it appears only with (recent) loans, accept for Gelmets *fura* ~ *fəra* "raspberry" (Tsakhur *xura*). While Tsakhur *x-* is the regular reflex of Proto-Lezgian **fʷ-*, *f-* is exceptional. Perhaps *fura* is an old loan from Lezgi *furu-fular* "berries". For Proto-Lezgian we have to assume a form **fʷər-* (Rutul *xur*, Tabasaran *urxar* (pl.tant.).

k: (m): *akːa* "door", *əkːas* "to love", *hikːin* "narrow";
g (i): *gurt* "shirt", *gandʒ* "rock", *gnej* "bread";
x (i,m,f): *xa* "skin", *xax* "handful", *xan* "ground", *juxan* "springtime", *o-oxanas* "to eat", *exalʲ* "at night";
xː (m): *qoxːaras* "to boil", *gʲixːes* "to lay down";
ɣ (i,m): *ɣalʲ* "mouth", *wiːɣaras* "to bear", *ɣarɣar* "sheep", *ɣam* "awn";
kʲ (i,m,f): *kʲira* "reward", *kʲidikʲ* "swelling", *kʲerkʲi* "mattock", *oˤkʲ* "yoke", *dakʲ* "father";
gʲ (i,m,f): *gʲaːzas* "to milk", *gʲaʔas* "sit down (sg.)", *gʲoharas* "to fall (from a tree)", *gʲəˤχaˤs* "to fall", *gʲeːd* "much", *gʲuˤχ* "defamation", *maˤgʲ* "hairdressing";
xʲ (i,m,f): *xʲan* "water", *xʲid* "dried manure", *dexʲe* "necessary", *t'ixʲ* "pimple", *wixʲ* "flea", *k'uxʲ* "beak";
xʷ (rare): *xʷar* "hole";
q (i,m,f): *qales* "to come", *quf* "nose", *ɢaqa* "outdoors", *məˤq* "worm", *aˤq* "schadow", *məˤq* "worm";
q' (i,m,f): *q'uq'* "egg", *q'om* "hight", *q'atfːi* "hat", *q'əf* "butter", *q'əʲdʲim* "winter", *jəq'* "broth", *amq'al* "turned around", *leˤq'* "eagle";
qː (rare): *zaqːa* "under me", *zaqːum* "bitterness", *tawaqːa* "request";
ɢ (i,m): *ɢarg* "ram", *ɢək* "consumables", *ɢul* "window", *ɢəˤlʲ* "summer", *manɢal* "heat";
χ (i,m,f): *χiw* "village", *χaˤm* "night", *diχ* "son", *alijχas* "to fly", *hoχas* "to run", *sirəχ* "seam";
χː (m): *heχːen* "too much", *qəzaχːas* "to meer", *laχːa* "stoven";
ʁ (i,m,f): *ʁu* "you (sg.)", *ʁurʁiː* "gravel", *ʁon ~ ɢon* "partridge", *jiʁ ~ jəʁ* "day", *nʲaʁ* "tear", *baʁ* "twig";
qʷ (rare): *nʲaqʷ* "straw", *maˤqʷ* "oak";
qʷ' (rare): *aqʷ'a* "face", *nʲaqʷ'* "earth", *q'arqʷ'* "lip";
χʷ (i,m,f): *χʷa* "dog", *meχʷ* "tale";
qˤ (i,m,f): *qˤaˤtqˤe* "reading", *qˤal* "badness", *qˤaˤn* "raven", *məˤqˤ* "oak[23]", *aˤqˤ* "schadow" (~ *aˤq*), *jəˤqˤ* "ashes";
qˤ' (i,m): *qˤ'əf* "blade", *qˤ'oˤ-* "two", *qˤ'aˤb* "tub", *hoˤqˤ'iˤn* "fat", *aˤqˤ'əˤn* "broad", *jəˤqˤ'* "back";
ɢˤ (rare): *ɢˤəˤje* "hare", *ɢˤəˤlʲ* "summer" (~ *ɢəˤlʲ*);
χˤ (i,m,f): *χˤeˤra* "big", *χˤoˤ* "flour", *χˤaw* "sky", *aˤrχˤ* "wool (of sheep in springtime)", *joχˤa ~ joχːˤa* "mill";
ʁˤ (i,f): *ʁˤaˤbref* "blackberry", *jəˤʁˤ* "bridge", *aˤʁˤaˤ* "hair comb";
ʔ (m,f): *haʔas* "to do", *tsʲ'eʔ* "goat", *maʔ* "bakon", *gʲuʔre* "have a sit (man)!";
h (i,m,f): *hagʷas* "to show", *hikːʲin* "small", *hoχas* "to run", *hor* "lies", *lahar* "home work", *maˤh* "brain";
hˤ (i): *hˤaf* "dirt", *hˤaˤnka* "wine-skin", *hˤeˤwaˤ* "lentil";
j (i,m,f): *jik'* "heart", *jatsʲ'* "stone", *aˤjne* "glas", *wejba* "needle", *abatːej* "father-in-law";
m (i,m,f): *mək* "ice", *maʔ* "fat", *maʁ* "ploughshare", *mətsʲ* "wind", *miz* "language, tongue", *ɢˤumaˤn* "warm", kuma "smoke", q'*əʲdʲim* "winter";
n (i,m,f) *nak ~ nʲak* "milk", *natsʲ'* "reed", *naʁ ~ nʲaʁ* "tear", *nisːe* "cheese", *t'enk'* "drop", *ok'anas* "to write", *k'on* "game (animals)", *k'ut'un* "kidney[24]";

[23] *məˤqˤ* or *məˤqˤʷ* is a free variant of *maˤqʷ* "oak".

l (i,m,f) *ləˤq'* "eagle", *lal* "dumb", *luk'un* "bark of the birch", *balkan* "horse", *dalkʷ* "container", *ɢul* "slave", *fal* "cloth", *gal* "the ara tree", *ɢəˤl* "summer[25]";
lʲ (i,m,f) *lʲiga* "front", *lʲamba* "fence", *kəlʲk* "remains of wool", *wuk'ulʲ* "head", *ɣalʲ* "mouth", *ɢulʲ* "window", *galʲ* "on us", *səkəlʲ* "fox";
r (m,f): *k'asrə* "50 pounds", *kʲira* "reward", *pəran* "dark", *werəʁ* "sun", *χˤeˤra* "big, old", *uxur* "barley-water", *wuhur* "narrow street", *paχər* "oxidization", *t'elʲamp'er* "musik instrument[26]".

The consonants given in brackets in (3) show that secondary correlations such as palatalization, labialization, and pharygealization play an important role in the system of Tsakhur consonants, though it is hard to tell to which extent this role is phonematic. In many cases labialization has to be treated as biphonematic (*C+w*), whereas pharyngealization of consonants may be the result of an assimilation to pharyngeal vowels. Palatalization, finally, is often conditioned by the adaptation of Russian articulation.

Archi is the only Lezgian language that shows a series of laterals. This series, common to all Awaro-Andian and (in part) to all Cezian languages, must also be reconstructed for Proto-Lezgian, it is, however, generally changed elsewhere (either to uvular or to velars). Tsakhur has the following reflexes of these laterals[27]:

PL *tɬ'	>	Tsakhur *k'* / Gelmets *q'*
PL *tɬː'	>	Tsakhur *k'* ~ *g* / Gelmets *q'*
PL *ɬ	>	Tsakhur *x*.

Except for the treatment of old laterals, Gelmets as well as the (sub)dialect of Sabunchi differ from the Tsakhur system of consonants only marginally. Gelmets has changed Proto-Tsakhur *ɢ to *ʁ* in all positions (cf. *ʁəʔje* "hare" (Tsakhur *ɢˤəˤje* "hare"), *ʁoˤn* "partridge" (Tsakhur *ɢon* ~ *ʁon*), *ʁijna* "today" (Tsakhur *ɢijna*) etc.). Sabunchi has retained the uvular stop *ɢ* in initial position and after *n*, elsewhere it shows the same change as Gelmets. Both (sub)dialects have a distinct phoneme *ʒ* (which merges with *dʒ* in Tsakhur proper except for loans): Gelmets: *həʒiwa* "cream", *gʲeʒe* "hive", *q'abaʒur* "carrot", *tʃoʒ* "brother"; Sabunchi: *ʒeʒaˤq'* "wasp", *ʒərʒəra* "dragon-fly".

The tendency to eliminate labialized phonemes, which can already be observed in Tsakhur, is much stronger in Sabunchi, though unpredictable. The same is true for the palatalization of final consonants in Gelmets: Although final palatals are much more preferred than plain consonants, it is hard to tell what conditioned this tendency (the vowel quality does not

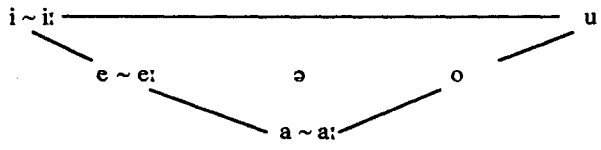
[24] With some lexemes, a palatalized nasal (*nʲ*) seems to be preferred. But there are no minimal pairs (*n* vs. *nʲ*) which would allow us to propose a phoneme *nʲ*.
[25] Final (non palatal) *-l* does violate an important sound law of Tsakhur: PL *-*l* > *-w* when preceded by a non palatal vowel. Else the reflex is *-lʲ* (also resulting from PL *-*li*). Hence, most of Tsakhur lexemes with final *-l* should be regarded as possible loans.
[26] PL **r-* is either dropped in Tsakhur or developed to *j-*. Some "Tsakhur looking" items like *rab* "awl", *rəxʲ* "saliva", and *roːtsʲ* "cradle" are loans (mostly from Rutul, some (like *roːtsʲ*) from Armenian. However, the case of *roːtsʲ* seems to be somewhat complicated, as we must assume that the reduction of arm. *ororoç* "cradle" > *roːtsʲ* took place at a time when initial *r-* still was allowed.).
[27] I only mention those laterals the reconstruction of which can be taken for granted.

play a role, as can be seen from examples like *wəsʲ* "buffalo", *nasəlʲ* "lineage", and *pəlʲ* "money").

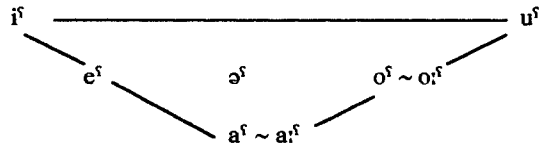
2.1.2 Vowels

The vowel chart of Tsakhur (proper) draws the following picture:

(5) Plain:



Phar.:



Tsakhur has developed a quite complex system of vowels, contrasting long from short, and plain from pharyngealized phonemes (nasals are not known).

In general it can be claimed that the basic system of vowels in the proto-language was very much like that of present Eastern Samur[28]. Most labialized vowels (except *y*) are secondary, often introduced through pressure from Azeri (*o* mainly resulted from *-*Cʷ*- complexes). Long vowels came in because of vowel contraction or because of the lengthening of vowels in stressed positions.

In Tsakhur, long vowels resulted especially from the contraction of vocalic class markers with surrounding vowels, though sometimes we can also find non secondary contrasts (mainly *a* ~ *aː*), e.g. *əkʼar* "illness" ~ *əkʼaːr* "being ill", *ma* "on" ~ *maː* "there", *ajχan* "moving" ~ *ajχaːn* "saw" (GEN). The main source for long vowels are complexes of the type Vw(V)(j) and Vy(V). (5) gives a list of the resulting phonemes:

(5)	awa	>	aː	owo	>	oː	iwo	>	oː
	iji	>	iː	awaj	>	aː	ij	>	iː
	awaj	>	aː	ej	>	eː (or iː[29])			
	ij	>	iː	aj	>	iː	oˤj	>	aːˤ
	uwuj	>	oː	ow	>	oː			

[28] Proto-Lezgian probably had **a*, **i*, **u*, and **y*. It must be said, however, that the system of vowels in Proto-Lezgian did not have that degree of functional load that was characteristic for the consonants. In this respect, Proto-Lezgian equaled very much present-day West Caucasian languages (cf. Schulze 1988a for details, as opposed to the views of Nikolayev/Starostin 1994, claiming that West Caucasian is a strongly innovating descendent of Proto-North Caucasian, which itself is thought to be very close to Proto-East Caucasian).

[29] *iː* only if **ej* was in a non initial syllable and preceded by *i*.

These processes which are of considerable importance regarding the emergence of ablaut phenomena will be exemplified in the appropriate chapter of morphology.

The question of the phonematic status of pharyngealized vowels still remains unsolved. In many cases we have to treat them as reflexes of surrounding pharyngealized consonants. However, sometimes these vowels appear with consonants, which clearly are non pharyngeals, e.g.:

(6)	aˤbəˤn	"cat"	(cf. *abən* "those, who are inside")
	oˤd	"bile"	
	uˤlʲdʒuˤm	"week"	
	tʃuˤmk	"whip"	
	juˤtʃʼuˤd	"nine"	
	oˤkʼ	"yoke"	

Hence it must be concluded, that pharyngealized vowels may have an independent status within the phonemic system. Perhaps this system is based on a (former) strategy to differentiate lexical items through suprasegmental features, which themselves had been realized by pharyngealization. This can also be seen from the fact that if pharyngealization appears in polysyllabic words this feature spreads over nearly all syllables, cf.

(7)	ʕʼaˤdaˤt	"habit, custom" (< arab. *ᶜadat* "dto.")
	qˤaˤdqˤəˤ	"having read"
	xˤiˤmaˤn	"wild"
	joxːˤaˤ	"mill"
	miˤtʃaˤd	"in the morning"
	jaˤqːeˤ	"on the road"

The feature [+phar] seems to start with the accentuated syllable and then spreads to the following syllables. That is why in some words, the vowel of the first syllable (mostly a preverb) is not pharyngealized, whereas all other vowels are, cf. *altʼáχːˤiˤj* "swelling", *bədéˤʁ* "gossip", *jiqʼˤɵˤn χew* "heavy burden" (from *jiqʼˤ* "heavy").

The (sub)dialect of Sabunchi tends to substitute pharyngealized vowels by palatals, a process that is well known especially in those SECL that are in closer contact with Azeri[30]. This tendency can be interpreted as a try to imitate the Azeri system of articulation without loosing the distinctiveness of the vowels in question, cf. (*aˤ* > *æ*, *oˤ* > *ø*, *uˤ* > *y*):

	Tsakhur	**Sabunchi**	
(8)	aˤjne	æjnæ	"glas"
	oˤdʒ	ødʒ	"badness"
	oˤruˤʃ	øryʃ	"driving in"
	tʃuˤmk	tʃymk	"whip"
	oˤkʼ	økʼ	"yoke"

[30] Also, this is a typical feature of what can be called "Young People's Tsakhur" (cf. 1.3).

The use of palatalized vowels is not restricted to old pharyngealized vowels. Sabunchi (and partly) Gelmets have extended this technique to plains vowels (even if they do not meet a palatal surrounding). Hence, palatalization can also be regarded to be at least partly a suprasegmental feature, which indicates an articulation mode "à la turque" rather than the emergence of true palatalized vowel phonemes.

2.1.3 Phonotactics

Tsakhur allows the following types of C/V-distribution within single syllable words[31]:

V:	*o:* "up", *a:ˤ* "going";
CV(:)	*do* "name", *sa* "one", *zə* "I", *xa:* "at home", *ʃa:* "there", *sʲo* "bear"; *tʲu* "spittle";
VC	*ad* "inside", *al* "devil", *eb* "blood", *etʃ'* "apple", *uˤm* "ring", *itʷ* "honey", *ulʲ* "eye", *ok'* "hay", *os* "log, chunk", *oˤkʲ* "yoke", *oˤʤ* "evil";
VRC	*urg* "lamb", *aˤrd* "fat-tail";
CVC	*dakʲ* "father", *ɢulʲ* "window", *ɢˤeˤl* "summer", *miz* "language", *mətʂ* "wind", *nʲak* "milk"; *waz* "moon", *jiz* "snow";
CVRC[32]	*dalkʷ* "sort of meal", *gan* "rock", *ɢarg* "wether", *q'arqʷ* "furs", *gurt* "shirt", *kəlʲk* "remnants of leather", *kʷ'artʃ* "manger", *t'enk'* "drip".

Another important phonotactic restriction is observed in connection with *r*. Whereas Eastern and Southern Samur as well as Rutul allow it in every position, it is changed in the other languages under specific conditions: Archi, Khinalug, Udi and (partly) Tsakhur do not allow *r* initially, changing it to *d-* (in Archi) or *z-* (under certain conditions in Khinalug), or dropping it completely (in Udi). In Tsakhur, there is a tendency to replace *r-* by *j-*, but contrary to all other SECL is does not allow *r* in (absolute) final position. (9) lists three word pairs in order to demonstrate the facts:

(9) Lezgi *rib*, Tabasaran *rib* ~ *riw*, Aghul *reb*, Rutul *rab*, Tsakhur *rab* (loan from Rutul), Kryz *ræb*, Budukh *ræb*, Archi *dab* "awl", Udi *eˤb-sun* "to sew";
Aghul *raq'*, Rutul *riq'*, Tsakhur *jəq'*, Archi *diq'*, Khinalug *ze*, Udi *opt'aj* "soup";
Lezgi *ruqˤ*, Tabasaran *ryq*, Aghul *ryq:*, Rutul *riqˤ*, Tsakhur *jiqˤ*, Kryz *ræqˤ*, Budukh *req*, Archi *diq'*, Khinalug *zek'*, Udi (Vartashen) *iq'* "ashes".

As it has been said above, final *-r* is quite weak in Tsakhur, being either dropped, changed to *-j*, or retained before an epithetic vowel (often *-a*). Some exceptions are *ar* "inside", *ʁarʁar* "sheep", *hor* "lies", *lahar* "homework", *makar* "slyness", *paχər* "oxidation", *sakar* "dam", *uχur* "barley-water", *wuhur* "narrow street", *χʷar* "hole", *zer* "cow", and *k'ar* "illness". Final *-l* behaves quite similar in Tsakhur: Generally, it is changed to *-w*, but it is retained before epithetic *-a* or when heavily palatalized by a (then dropped) final **-i* (> *-lʲ*).

[31] Cf. Ibragimov 1981 for details (his data stem from the dialect of Gelmets and slightly differ from those given here).
[32] In loans, Tsakhur also allows final clusters without sonants, cf. *mast* "sheep milk" (geo. *mast'ao* "a meal made out of butter-milk and rice" (?)), *rəʂ'q'* "food" (aw. *rizq:i* "food"), *dost* "friend" (Persian).

2.1.4 Phonetic processes

There is a tendency towards vowel harmony in Tsakhur, though it is not as systematized as e.g. in Turkish. The language shares this feature especially with Rutul and the Eastern Samur languages, it is less observed in Southern Samur, Archi, and Udi. Vowel harmony mainly effects suffixes, e.g. the class marker *-ɣd* (class IV): it appears as *-d* after non labial vowels, as *-ud* after labials, and as *-Vˤd* after pharyngealized vowels, cf.

(10) Tsakhur: *χʲeb-əd* "three", *jiχ-əd* "six", *joq'-ud* "four", *molʲ-ud* "seven".

The vowels of the plural morpheme *-Vr*, however, depends on the quality of the foregoing consonant (*-er* goes with palatalized, *-ar* with non palatalized consonants). Final pharyngealized phonemes prefer *-aˤr*. Yet, It has to be stressed that vowel harmony is still a quite unpredictable tendency in the language, and in many cases it depends on the idiosyncrasy of the speakers.

Another important process is the labialization of consonants due to the loss of unstressed labial vowels following them. This process can be observed in all SECL that have labialized consonants, and in many cases it is even not bound by stress conditions. (11) lists some examples from Tsakhur:

(11) Tsakhur: *balʁʷan* "fish-" (cf. *baluʁ* "fish"), *ʧarχʷan* (genitive of *ʧaruχ* "bast-shoes"), *qeq:ʷas* "to dry" (< **qeq:əwas*), *qidekʷ'ij* "immortality"(< **qidek'wij*), *hekʷ'as* "to bake" (< **hek'was*), *hadʁʷas* "to tear" (< **hadʁəwas*).

Another process often described for the SECL is that of assimilation. Especially clusters like *-mb-*, *-nb-*, *-r+C-*, *-l+C-*, *-n+C-*, and so on tend to be homorganized on the basis of the second consonant. Examples from Tsakhur are:

(12) Tsakhur: *ʤammə* < **ʤam-bə* "meals", *q'ommu* < **qom-bu* "peaks", *k'ən-na* < **k'əl-na* "small, class I-III), *semmə* < **sen-bə* "years".

Finally, a quite typical feature of all SECL is that of metathesis. In many cases it seems that (especially with loans) the order of phonemes within a word is relatively free. There are not yet any detailed studies on this topic available, but we can observe the tendency to place homorganic consonants into a closer contact or to avoid -CR-structures, cf. Tsakhur:

(13) *t'aq'ra* > *q'at'ra* > *q'art'a* "drip", *qelezəm* > *q'eleməz* "brine", *ɢarʁudali* > *ɢardurali* "corn", *k'əbərχ* > *k'ərbəχ* "long-eared", *sək'əlʲ* > *səlʲək'* "rye".

2.1.5 Prosodic features

Tsakhur, as most other (Nothern) SECL has a dynamic word accent[33], which falls on the second syllable of polysyllabic words[34]. Suffixes normally are unstressed (with the exception of the plural morpheme *-Vr* and *-es* encoding the future tense[35]), cf.

(14) *k'umk'úm* "saucepan", LOC *k'umk'úm-e*
balkán "horse", DAT *balkán-ɔs*, but *balkan-ár* "horses"
akɜ́ "dooryard", LOC *akɜ́-la*

Verbs are always stressed on the first syllable (with the exception of the future tense), which is also true for monosyllabic nouns of the type CVC or CVRC, cf.

(15) *sos* "daughter-in-law", ERG *sós:-e*
q'om "peak", LOC *q'óm-alʲ*
boz "grey", GEN *bóz-un* ~ *bóz-na*

However, monosyllabic words either starting with a vowel (VC or VRC) or which are open (CV) are stressed on the (first) suffix in inflection, cf.

(16) *ulʲ* "eye", LOC *ulʲ-él, ulʲ-él-χa*
urɣ "lamb", PL *urɣú-bə*
ʦ'a "fire", DAT *ʦ'aj-ís*

As had been said above (2.1.2) labialization, palatalization, and pharygealization can at least partly be regarded as prosodic features in Tsakhur. This procedure which is typical for some other SECL is not yet well enough understood. However, there is clear evidence that in some cases lexical items are placed under a "pharyngealized" or "labialized" tone to contrast them with words in plain pronunciation. Moreover, this "tone" clearly depends on the pace of word stress: Only in vary rare instances the above mentioned secondary articulations start before a stressed syllable.

2.1.6 Morphophonemics

Morphophonemic features play an important role in Tsakhur, especially in verb inflection. Ablaut phenomena met here are of relatively recent origin (resulting from the contraction of

[33] This stress pattern is contrasted e.g. in Udi by a strong tendency towards an ultima accent (which can also be found in Kryz and Budukh). Thus Tsakhur behaves much more like the typical East Caucasian languages than those more southerly languages. - Kodzasov/Murav'eva 1982 have described tonal features (four tones) for Tabasaran, yet they seem not to have any functional value. During my fieldwork I never noted such a discriminating function of tones in SECL (nor do mention them the written sources, as far as I can see). In Tsakhur, stressed syllables tend to have a high pitch, which is eventually reserved for long vowels in Gelmets.

[34] Stress does not depend on the vowel quantity in Tsakhur. Thus short and long vowels can both be met in (un)stressed position.

[35] In fact, it is the theme vowel *-e-* which is stressed in the future tense forms, cf. below.

sonants and surrounding vowels. They shall be discussed in the appropriate chapter (cf. 4.7.2). However, in some cases (especially in noun inflection and in TAM formation) an older type of ablaut can be recorded (cf. ABS *deχ* ~ *diχ* "son", ERG *duχe:* ~ *diχawe*, GER.PRES *-a* vs. GER.PAST *-u* etc.). However, alternation of stem vowels is much rarer in Tsakhur than in many other ECL. In general, the nominal ablaut in SECL still needs an explanation.

3. Word formation

The word formation of nouns in Tsakhur follows the general pattern of word formation in the Samur languages: Both suffixation and composition can be met. The following suffixes are more or less productive:

-aj ~ -aˤj ~ ej ~ ij: A quite productive morpheme used for ethnonyms, family names, relatives, nomina agentis, nomina instrumenti etc. Very often the basic lexeme is no longer used, cf. *jix-ij* "a Tsakhur speaker", *mutur-ij* "Georgian", *waz-aj* "Wazaj" (fem. P.N., cf. *waz* "moon"), *t'ik'-ij* "a little wine-skin" (cf. arm. *tik* "wine-skin"), *sʲurtʲul-ij* "gullet", *did-ej* "grand-mother", *ab-aj* "mother-in-law". The prototypical meaning of this suffix perhaps simply is relational, in some cases such as *adam-ij* "man" it seems to have a singulative function.

-lla ~-walla: Abstract nouns from adjectives: *jig-walla* "goodness", *mek'-walla* "youth", *kasib-walla* "poverty", *bat'ra:-lla* "beauty", *cˤuma:ˤ-lla* "heat".

-ulaj ~laj: Nouns referring to specific properties of the basic lexeme: *waˤqˤ'aˤ-laj* "simple minded" (*waˤqˤ'aˤ-* "sheep"), *jabə-laj* "devoted" (cf. *jabə* "gelding"), *q'otʃ'e-laj* "gossip" (cf. *q'otʃ'e* "leathern bag").

-at'aj ~ t'aj: Same as *-laj* with a negative connotation *wuxun-t'aj* "glutton" (cf. *wuxun* "stomach"), *nʲaʁ-ət'aj* "quarrelsome" (cf. *nʲaʁ* "tear").

-ebaj ~ abaj ~ baj: Denominals referring to some characteristic features of the basic lexeme: *ʦʲoj-ebaj* "burdened" (cf. *sʲo* < **sʲoj* "bear"), *sile-baj* "biting" (cf. *sili* "tooth"), *ɣalʲ-baj* "gaper" (cf. *ɣalʲ* "mouth").

-qan ~ qˤaˤn ~ aqan: Nomina agentis (like the loans suffix *-ʧi* (< az. *či*): *joˤxˤraˤ-qˤaˤn* "miller"[36], *gew-aqan* "shepherd" (cf. *gew* "sheep").

-alij ~ -lij: Nomina agentis: *cˤaˤn-alij* "shepherd caring for milk sheep".

[36] This is the form as given by Ibragimov 1990:84. Normally, "mill" is *jox:ˤaˤ*, and *joχ:ˤaˤ-qˤaˤn* for "miller" has been confirmed by my informant.

In many instances, case morphemes (genitive and locatives) are uses as word formation suffixes. Moreover, a number of loan suffixes (esp. from Azeri) has entered the Tsakhur paradigm (among others: *-lə* (nomina originis), *-səz* (α-privativum), *.-dar* (pers., only in loans). Composition involves two nouns, the first of which is either in the absolutive (rare) or in the genitive. Examples for genitives are *ulena-xˤaˤne* "pupillary" (eye.GEN-star), *k'ərajn-xeb* "temple" (ear.GEN-fundament), *wuxnen-dan* "navel" (stomach.GEN-cheek)[37].

Finally, reduplication plays an important role in word formation. Either full reduplication appears, or only parts of the lexeme are reiterated (sometimes changed phonetically). Full reduplication is e.g. given in *t'a-t'a:* "stick", *k'um-k'um* "dish", *ʁar-ʁar* "empty nut", *tʃ'iw-tʃ'iw* "song-bird". Modifications are *k'elʲ-k'am* "liver", *zəl-zam* "milt", *q'an-q'aj* "raven". Prereduplication appears e.g. in *sir-sifat* "a kind", *zir-zibilʲ* "fat", *p'aˤp'al* "bark of a fruit".

Verbal word formation is based on two principles: First, the metaphorical extension of preverbial forms (cf. 4.7.3), and second, nouns incorporation, cf. *iʃ-haʔas* "to work" (work-to make), *uba-haʔas* "to kiss" (*uba* "kiss"), *miz-əˤwχˤaˤs* "to lick" (*miz* "tongue").

4. Inflectional Morphology

4.1 Noun Classes

Except for Udi, Aghul, Lezgi, and Southern Tabasaran all SECL know a system of morphologically marked noun classification. By this is meant that nominal as well as pronominal referents are classified according to nowadays somewhat obscured semantic categories. The morphological means to encode this classification form the basis for agreement phenomena in the single languages. As the system of noun classification is one of the basic aspects of the morphosyntactic and morphosemantic operating systems of the languages, it will discussed in more detail here than the other linguistic categories.

Prototypically, all SECL in question operate on a four class system[38], which is reduced only in Nothern Tabasaran (leading to a system of two classes). The semantic allocation of nouns with respect to the different classes today is highly formalized, though some specific features of the single classes can still be enumerated. The first class (CL I) includes only human males, sometimes also the names for superficial beings, cf. for Tsakhur:

(17) CL I: *adamij* "man", *wəɣəlʲ* "male being", *tʃodʒ* "brother", *deχ* "son", *dakʲ* "father", *didej* "grandfather", *gade* "youngster", *ɣenif* "god", *idaːg* "prophet", *malʲaʔik* "angle".

CL II includes all female human beings, as well as some mythic names, cf.:

(18) CL II: *xunaʃːe* "woman", *jedʲ* "mother", *adej* "grand-mother", *mamasij* "sister of father", *itʃij* "girl", *jitʃi* "sister", *jif* "daughter", *sos* "daughter-in-law", *qˤ'aˤj* "concubine", *tʃengij* "prostitute", *magadʒin* "witch", *tsitsij* "doll".

[37] This type appears especially for body part terms.

[38] The system of noun classification in the SECL is discussed e.g. by Alipulatov 1974, Xajdakaov 1980, Magometov 1980, Drossard 1982, K'aχaʒe 1984, Schulze 1988b, Schulze 1992, Schulze(-Fürhoff) 1992, and Schulze (in preparation).

Whereas the distribution within these two classes is quite predictable (at least in the SECL)[39], the two remaining classes are far from having a clear semantic specification. Both classes include animals as well as different types of objects and abstract nouns. However it has been shown that the two classes can be seen as the poles of a (hierarchic) parameter, which is prototypically defined by the following cluster of features: {[±important];[±large];[±socially relevant];[±active]}. The more these features are activated within the (connotational) semantics of a noun, the more likely is its inclusion in CL III[40]. This prototypical system together with its radial extensions can be retrieved from many subparadigms within the vocabularies of the SECL, cf. e.g. the distribution of domestic animals and their young ones in Archi:

(19)

III		IV	
xˤon	cow"	biʃ	"calf"
noˤʃ	horse"	uri	"foal"
tʃ'an	sheep"	tɬ:'al	"lamb"
ts'aj	goat"	motol	"kid"
gʷatʃi	dog"	k'onts'ol	"whelp"
gatu	cat"	gatulin lo	"kitten"
x:enne-heleku	hen"	helekulin lo	"chicken"

In Tsakhur, this parameter no longer functions on a synchronic level. (20) lists some words for each class in order to show the degree of formalization that the paradigms have already attained (CL IV includes only a small number of animals):

(20) CL III: *q'arg* "wether", *ɢˤəˤje* "hare", *dadal* "cock", *χotʃe* "snake", *ulʲ* "eye", *gəməlʲ* "rheumatism", also "cloud", *χiw* "village", *t'uʃ'* "twig", *nawur* "lake", *sifat* "face", *dʒik'rə* "tail", *mets* "wind", *maˤmaˤʁˤ* "wealth", *wuk'ulʲ* "head", *əˤməˤr* "life".

CL IV: *baluʁ* "fish", *aslan* "lion", *aˤrdawa* "drake", *χəlʲ* "hand, arm", *jik'* "heart", *ts'a* "fire", *jiz* "snow", *xʲan* "water", *jiwa* "iron", *kalle* "head", *bəˤt* "tail", *jaʃajif* "life".

The assumption of four classes in Tsakhur as well as in the other SECL in question is based on a paradigm of agreement morphemes, which will be discussed below. However, it has to be stressed that many languages including Tsakhur use specific paradigms to encode quite different types of semantic subcategorization. Whereas the feature of class differentiation can be described as being basic for agreement morphemes, some morphosyntactic paradigms such as plural and case marking, attributive inflection and so on show a secondary tendency to use the

[39] In Nothern Tabasaran these two classes are unified, now including all human beings (in the singular).

[40] A semantic motivation for class III and IV has first been proposed by Mel'nikov/Kurbanov 1964 (for Tsakhur). However, their is system seems quite vague (they propose a parameter of "importance") and shall not be discussed here in detail. Kibrik et al. 1977:55-66 have tried to show that this parameter works at least partly also in Archi (though the distribution of nouns in Archi is influenced much more by secondary processes than e.g. that of Tsakhur). In Schulze 1992 I have argued that the overall principle of noun classification in the East Caucasian languages is the allocation of nouns according to the question to which degree their referents are thought to exert a controlled activity in society (cf. also Klimov 1977:29 who treats the distribution of CL III and CL IV in terms of his "active typology").

inherent semantic aspects for class differentiation. In Tsakhur this is especially true for the ergative case and for adjectives. Here, we can meet e.g. the following types of (overt) noun classification[41]:

(21)

	Agreement	ERG	Adjectives	OBL SA[42]
CL I	a	a	a	a
CL II	a	a	a	b
CL III	b	b	a	c
CL IV	c	b	c	c

Thus, in case marking there is a strong tendency to separate male human beings from the rest of the referents, whereas with attributive adjectives most of the animates are subsumed in one class. This technique is probably old (and sometimes even reestablished in SECL) and can only be explained by assuming a period when the basic classificational procedures were still vital (when a specific subparadigm induces noun classification, this will be indicated in the following chapters).

The interaction singular and plural noun classes plays an important role in Tsakhur as well as in every SECL. Basically, the singular morphemes are used to encode plurals as well, referring to specific aspects of the singular class encoded by the morpheme. Plurals are less distinctive than singulars that refer to discrete entities. Thus we have a merger of classes in the plural by reducing the distinctiveness of the semantic features in question. Human and males and females are fused in one class ([αsex;+hum]) and are encoded with the help of the singular class morpheme, that is next to class I and II in the hierarchy, i.e. by class III. The same holds for the plurals of class III and IV which merge into one class encoded by the singular class IV morpheme. (22) indicates the facts schematically:

(22) Singular: I < II < III < IV

Plural: [I+II] ← [III+IV] ←

Because there is a decisive clear cut between the human classes I and II and the rest of the referents, the class markers used as plural morphemes thus establish a simple dichotomy, separating human beings from the other referents. The basic formula for the prototypical distribution of class markers in given in (23):

(23)

	SG	PL
I	a	c
II	b	c
III	c	d
IV	d	d

[41] The system here proposed is by no means complete. There are many subtypes to be observed in nearly every SECL. "a", "b", "c", and "d" symbolize any difference in the inventories of morphemes used for class marking, they do not indicate *real* class markers.

[42] Oblique stem augments of nominalized adjectives (singular): CL I *-G^{w}-*, CL II *-G-*, CL III/IV *-ʧi-*, see below.

Tsakhur - as the other SECL - has varied this formula considerably. Historically, the single classes had been encoded by means of one morpheme each, cf. (24) which gives the proposed reconstruction of those morphemes for Proto-Lezgian:

(24)

	SG	PL
I	*w	*b
II	*r/j	*b
III	*b	*d
IV	*d	*d

The technique of class marking based on the system in (24) which is generally covert in the SECL: the noun that triggers class marking is not marked itself[43]. The basic strategy for agreement is to mark those constituents in a phrase that are thought to stand in a close or central relation with a head noun. Outside the noun phrases agreement plays a considerable role with respect to syntactic functions. Normally, a verbs goes with a noun in subject function (mostly marked by the absolutive, see below). However, in some languages, an ergative marked noun can trigger class marking even outside the NP (e.g. on adverbs, dative pronouns and so on). In this case, we have to assume that class agreement does not depend on the subject function of a noun exclusively, but can be dominated by certain pragmatic features (standard topics, attention flow etc.).

The Tsakhur inventory of class markers is given in (25), the list is neutral with respect to distributional aspects:

(25)

CL I:	r, j, w, ∅, e, a, u, -na, -da
CL II:	r, j, ∅, e, ə, -na, -da
CL III:	w, b, aː, eː, -na, -da
CL IV:	d, l, j, i, ∅, -n
CL I/II.PL	d, b, w, aː, eː, o, oː, ∅
CL III/IV.PL	d, j, w, i, aː, ∅

The distribution of these morphemes together with informations on their degree of semantic distinctiveness is given in the appropriate chapters (see adjectives, pronouns, and verbs).

[43] There are, however, some important exceptions: As many other ECL, the Lezgian languages have used class markers to identify the sex e.g. of brothers/sisters, boys/girls and so on (cf. Archi *ufdu* "brother" (< **w-əf-du-w*) vs. *dofdur* "sister" (< **r-əfdu-r*); Khinalug *fi* "son" (< **w-ifii*) vs. *r-ifii* "daughter" (Lezgi *rif*, Tabasaran *rif*, Aghul *ruf*, Rutul *rəf*, Tsakhur *jəf*, Kryz and Budukh *rif*, Archi *difdur* ("new-born daughter")). The system is much better preserved in the Nothern ECL and is based on old adjectives ("belonging to the same family", "young" etc.) - In Tsakhur nouns can be class marked secondarily when focused. But this is not a case of overt classification as the noun itself does not trigger the class marker, cf. 5.4.

4.2 Nouns

4.2.1 Number

Like Tsakhur, all SECL have a simple dichotomy, separating a plural from the singular. In some languages (e.g. in Udi) a tendency to encode collectives separately can be encountered. There is no clear evidence for a dual neither synchronically nor diachronically.

The SECL do not have means to indicate a nominal "singulative" morphologically, the singular is unmarked (at least in the absolutive). However, nominalized adjectives sometimes are marked with such a morpheme (see below). The plural morpheme itself is heavily allomorphic in many SECL (except the Eastern Samur languages)[44]. The reason for this can be a semantic classification (e.g. in Rutul), phonological or phonotactic restrictions, or simply a lexical distribution without any clear motivation. It has to be born in mind that many plural morphemes or allomorphs are restricted to the absolutive case. Plural marking with oblique cases can further complicate the paradigms. In these cases the languages in question show a considerably strong tendency away from agglutination towards a polyfunctional marking technique. Tsakhur plurals operate on the following set of morphemes/allomorphs:

(26) -bə, -bu, -ba, -er, -ar, -r, [-je, -apːə, -abar][45].

The distribution of these elements is essentially based on the syllabic structure of nouns: *-bə* and *-bu* are used with nouns ending in open syllables, *-er* and *-ar* follow finals consonants. However, there are as many exceptions to these conditions as instances which match them, so that the given formulation must be taken with care. (27) lists a couple of examples:

(27)

Plural **-bV**:			Plural **-Vr**:	
akːa-bə	"dooryards"		miz-ar	"languages"
χotʃe-bə	"snakes"		balkan-ar	"horses"
do-bə	"names"		etʃ-er	"apples"
sʲo-bu	"bears"		k'uk'-ar	"spoons"
dam-mə	"cheeks"	(< *dam-bə)	χiw-ar	"villages"
k'ut'um-bə	"kidneys"	(< *k'ut'um-bə)	jiw-ar	"trees"
jats̱-bə	"bulls"			
t'et'-bə	"colors"			
zer-bə	cows"		etc.	

The system of plural marking in Tsakhur comes close to that of Rutul, but this language shows, that the distribution of *-bV(r)* vs. *-Vr* can be based on pure semantic reasons: Here the morphemes with labials are restricted to inanimates, and *-Vr* to animates. It is hard to decide,

[44] Topuria 1973, 1995 gives detailed information on the formation of plurals in the ECL.

[45] The morphemes in brackets have a very limited distribution (*-apːə* appears with paired body parts (*ulʲ-apːə* "eyes", *xul-apːə* "hands"), *-ba* ~ *-abar* is restricted to kinship terms (*ʎfu-ba* "brothers", *jif-ba* "daughters", *diχ-ba* "sons", *xən-abar* "second born brothers/sisters")).

whether the semantic systematization in Rutul is old or an innovation[46]. The morphemes themselves probably are of Proto-Lezgian origin, they can (at least) partly be traced in all SECL.

4.2.2 Case System

All case systems in the SECL are mainly based on three principles: First, they are basically ergative, opposing an ergative case to the absolutive. Second, a considerable number of nouns in inflected on the basis of a "stem augment" (SA), which sometimes is identical with the ergative case. Third, there is a relatively strong opposition between "functional" and "local" cases[47].

Contrary to most other ECL, nouns in Tsakhur can be secondarily marked by class markers irrespective the case form in which they appear. The only restriction is that the noun must be in focus function. The class marked on the noun in question is determined by the syntactic role "subject" (hence nominal class agreement goes along with verbal class agreement, cf. 5.4). The general scheme would be:

(28) NOUN_{FOC}-CASE-CM $\text{NOUN}_{\text{SUBJ}}$ CM-VERB

If the noun in subject function of an intransitive sentence is topicalized itself, the resulting structure is somewhat alike an overt class marking technique (though secondary), cf. a sentence like

(29) deχ-∅-ər qar-ə
son(I)$_{\text{FOC}}$-ABS-I return-PAST
"The BOY returned."

As for the functional cases, all SECL know an (unmarked) absolutive, an ergative, a genitive, and a dative. Contrary to Tsakhur the genitive is subcategorized according to the feature of (in)alienability in Kryz, Budukh, and (partly) in Khinalug, there are two functional relevant forms of the dative in Udi. The paradigm of case inflection in Tsakhur comes close to the Lezgian prototype, it has, however, reduced the system of stem augments considerably[48]. Here a nominal SA *-n-* appears for nouns[-hum,+sg], cf. *nʲuq'-n-* (OBL) "soil", *nʲuq-n-* (OBL) "straw", *nʲik'-n-* (OBL) "dream", *nʲik-n-* (OBL) "milk". It is quite remarkable, but not yet explained, that in Tsakhur the SA *-n-* nearly exclusively appears with nouns of the structure *nVC[+velar/uvular stop]*. Nouns[+hum.+sg] have *-aj-* as a stem augment. Normally, SA starts with the ergative case, however, the SA *-aj-* may be excluded from this case, cf.:

[46] Cf. Ibragimov 1978:46-51 for details.

[47] A quite recent overview on case inflection in the ECL is given by Topuria 1995.

[48] The system of stem augments in the SECL is somewhat simplified in Eastern Samur, Khinalug, and Udi, but highly allomorphic in Rutul, the Southern Samur languages, and Archi. The distribution of stem augments probably was semantically motivated in Proto-Lezgian (as it is still in the Nothern ECL). However, the different means (perhaps originally based on an opposition [±male(human)]) has been very much obscured in the single languages, so that nowadays it seems more appropriate to consider them lexicalized.

(30)	ABS	ERG	OBL	
	deχ ~ diχ	duχ-e:	duχ-aj-	"son"
	nʲaqʷ'	nuqʷ'-n-en	nuqʷ'-n-	"soil"

The ergative case in Tsakhur is semantically sensitive, it operates on the dichotomy [±human] (= class I/II vs. III/IV). Nouns [+human] are encoded by means of the morpheme *-e ~ -e:*, nouns [-human] are marked by *-n ~ -an ~ -ən*, cf.:

(31)		ABS	ERG	
	I	tʃodʒ	tʃodʒ-e	"brother"
	II	jiʃ	jiʃ:-e	"daughter"
	III	χotʃe	χotʃe-n	"snake"
	IV	aslan	aslan-an	"lion"

In the plural the ergative marker *-e* is used exclusively. The genitive morpheme is highly functional in Tsakhur. It is sensitive for noun classes as well as for the syntactic role of the head noun. Regarding class we have a distinction between class I-III (*-(V)na*) and class IV (*-(V)n*) (the trigger is the (singular) head noun, if it is not inflected)[50]. If the head noun is inflected, the genitive morpheme is *-(V)ni*, regardless the class of the head noun. (29) gives the basic paradigm:

(32)	ABS	GEN	[+NP$_{ABS}$]	GEN	[+NP$_{OBL}$]
	tʃodʒ	tʃodʒ-na	[+ NP I]	tʃodʒ-ni	[+ NP I]
	"brother"	tʃodʒ-na	[+ NP II]	tʃodʒ-ni	[+ NP II]
		tʃodʒ-na	[+ NP III]	tʃodʒ-ni	[+ NP III]
		tʃodʒ-un	[+ NP IV]	tʃodʒ-ni	[+ NP IV]

The following examples demonstrate the distribution of the different case forms:

(33) q'urban-na tʃodʒ tʃ'alag-a ∅-əxa-wu
Qurban-GEN(I) brother(I).ABS wood-LOC I-be-PAST.CAUS
"Because Qurban's brother was in the woods..."

q'urban-na jiʃ tʃ'alag-a j-ixa-wu
Qurban-GEN(II) sister(II).ABS wood-LOC II-be-PAST.CAUS
"Because Qurban's sister was in the woods..."

[49] *deχ* "son" has an alternative ERG *diχ-aw-e* (perhaps from **diχ-aj-e:*).
[50] There is a strong tendency in Tsakhur towards a formal and functional syncretism effecting the non-human ergative (*-(V)n*) and the genitive of class IV heads (*-(V)n*). This type of syncretism is often observed in ECL, especially in Lak, Dargwa, and the in pronominal systems of SECL, cf. Schulze 1997b for details. At least historically we can assume that the opposition of *-e*-marked and *-n*-marked ergatives mapping the dichotomy [+sg,±human] was related to two different degrees of agentivity. Whereas a human agent was thought to have a high and autonomous control of a transitive action, non-human agents were seen much more in the light of a (natural) possessive structure (see 5.2.2).

q'urban-ni tʃodʒ-us kitab hi-wu
Qurban-GEN$_{OBL}$ brother(I)-DAT book(IV).ABS IV.give-PAST.CAUS
"Because he gave Qurban's brother the book..."

The dative (*-(V)s*) is one of the most consistent case morphemes in all SECL. Thus we have *dakːi-s* "to the father", *duχ-aj-s* "to the son", *jiʃː-es* "to the daughter" etc. Just like in Tsakhur, it also forms the verbal infinitive with a quite strong telic meaning in most of the other SECL. Table (34) summarizes the case forms in the singular discussed so far:

(34)			
	ABS	-∅	
	ERG	-e(:)	[+hum]
		-(V)n	[-hum]
	GEN	-(V)na	[head noun ABS, class I-III]
		-(V)n	[head noun ABS, class IV, all ABS plurals]
		-(V)ni	[head noun oblique]
	DAT	-(V)s	

In the plural case forms are added to the plural morpheme, which has a supplementary stem augment *-ʃ-* in the oblique cases. To *-ʃ-*, a foregoing *-r-* (plural) is assimilated, leading to the lengthening of a preceding vowel, cf.

(35)		SG	PL	
	ABS	tʃodʒ	tʃu-ba[51]	"brother"
	ERG	tʃodʒ-e	tʃu-bi-ʃ-e	
	GEN	tʃodʒ-na/-un/-ni	tʃu-bi-ʃ-da/-in/-di[52]	
	DAT	tʃodʒ-us	tʃu-bi-ʃ-is	
		SG	PL	
	ABS	balkan[53]	balkan-ar	"horse"
	ERG	balkan-an	balkan-aː-ʃː-e	
	GEN	balkan-ana/-an/-ani	balkan-aː-ʃː-ina/-an/-ani	
	DAT	balkan-us	balkan-aː-ʃː-is	

The system of local cases in Tsakhur is no longer as systematic as for instance in Lezgi or Tabasaran. These languages exhibit quite refined paradigm which iconically differentiates the relational setting of a trajector with respect to its landmark (called "series") from the type of movement that is exerted with respect to the landmark (called "case"). The local cases are

[51] The drop of *-dʒ* in the plural forms is exceptional. The opposition of *-ba* (PL.ABS) and *-bi-* (PL.OBL) shows that in some cases the ABS is morphologically marked in Tsakhur.
[52] The genitives *tʃubiʃda* and *tʃubiʃdi* show the replacement of *-na/-ni* by *-da/-di* after palatal fricatives.
[53] This word is often thought to have a Proto-East Caucasian background (cf. Nikolayev/Starostin 1994:285 < **bʃaltkːe* "a big hoofed animal"), cf. lezg. *balk'an*, kryz. *balkæn* "horse". However, we might think of it as being a loan perhaps from Arabic (via Osmanic and/or Persian), cf. arab. *ablaq* "pied", pers. *balaɣ* "pied horse", osm. *eblaq* "pied horse", *belqaː* "mare with white trammels", *balaq* "pied". *balkan* could then stem from a Persian plural *balaɣ-an*.

marked by the interaction of these two aspects. (36) gives a quite rough formula of the coding strategy:

(36)

Series	On	Under	In	In front of	Behind
Case	-a-	-b-	-c-	-d-	-e-
ESS -x	-ax	-bx	-cx	-dx	-ex
ALL -y	-ay	-by	-cy	-dy	-ey
ABL -z	-az	-bz	-cz	-dz	-ez

In Tsakhur this has been reduced to basically four local distinctions: Subessive, Superessive, Adessiv, and Inessive, cf.

(37)

Series / Case		SUB -Vkʲ-	SUPER -Vlʲ-	AD -Vs-	IN[54] -eː/-aː
ESS	-∅	-Vkʲ	-Vlʲ	[-Vs]	-eː/-aː
ALL	-qa	-Vqːa	-V-l-qa	-Vs-qa	-eː-qa/-aː-qa
ABL	-(ntʃ)e	-kʲ-e	-Vlʲ-e	-Vsː-e	-eː-ntʃ-e -aː-ntʃ-e

The basic semantics of these complex forms are still vital in Tsakhur, though they are sometimes hard to discriminate. Most of the case forms can also be class marked, cf. 5.2. and the following examples:

(38) tʃodʒ daxa-l-e-r ɢiʔitʃ'u ark'ən-na
brother(I).ABS roof$_{FOC}$-SUPER-ABL-I I.descend.PAST I.go-away-PAST.GEN
"Having descended from the ROOF the brother went away."

tʃodʒ-usː-e-b balkan alʲa-b-t'u
brother$_{FOC}$-ABL-III horse(III).ABS PV-III-take.PAST
"He took the horse from (his) BROTHER."

ʦaj-eː-qa-d kabab-əs tʃura gʲixː-ə
fire$_{FOC}$-ALL-IV shashlik-DAT meat(IV).ABS I/II(PL).lay-PAST
"They laid the meat for the shashlik into the FIRE."

The trigger for class coordination always is the absolutive marked noun. In these cases class marking has focus function (basically, the "Daghestanian type" of CM being used). There is, however, a second paradigm of local cases, in which the markers of the series are followed by the class sensitive genitive morphemes. In these cases, the trigger noun is clearly definite. The following morphemes are used:

(39)

-kʲ-na/-in/-ni	Subessive
-lʲ-na/-an/-ni	Superessive

[54] Only used with nouns [-anim].
[55] The morpheme *-qːa* resulted from a contraction of the group *-*kʲqa*.

-l-qa-na/-n/-ni	Superessive/Ablative
-na-na/-n/-ni	Adessiv (membership)[56]
-s-ana-na/-n/-ni	Adessiv (location)[57]
-s-da/-in/-di	Dative (also used as a comparative)
-s-qa-na/-n/-ni	Adessive/Ablative

Cf. as an example:

(40) jorʁa balkan-nana adamij qˤ'aˤs qeqe-na deʃ
ambling horse-GEN.AD-I man old become.FUT-I NEG
"The man who has an ambler does not age."

jug-na balkan-na-n χaw məs qeqe-n deʃ
good-III horse-GEN.AD-IV family(IV).ABS hunger become.FUT-IV NEG
"The family (lit. "home") which has a good horse, will not suffer from hunger."

There are another two cases which structurally belong to the paradigm of local cases which, however, own a quite specific semantics: *-Vkʷa* is used as a comitative, whereas *-Vk'-le* encodes certain experiencers in so-called affective constructions, cf.:

(42) dakː-ikʷa-r jedʲ idak'ən-na
father-COM-II mother(II).ABS$_{FOC}$ II.NEG.go-out-PAST.GEN
"MOTHER did not go out with (our) father."

jed-ik'le urus miz w-aʦ'a wo-b
mother-AFF Russian language(III).ABS III-know.PRES AUX.PRES-III
"Mother knows Russian."

4.3 Adjectives

As had been said before, adjectives are not inflected as attributes in many SECL. Some languages still allow class agreement, especially Western Samur and Archi. When nominalized the adjectives often follow the inflectional paradigm of nouns, however, there are considerable peculiarities to be observed. Tsakhur itself shows one of the most complex system of adjective inflection in all SECL. The attributive of a nominal head agrees with its head according to different paradigms. The most productive system is based on the above mentioned genitive morphemes, it classifies nouns according to the inferential feature [±important]. The distribution is given in (43):

(43)

SG			PL	
I	a	[+important]	d	[αimportant]
II	a	[+important]	d	[αimportant]

[56] Probably a doubled genitive.
[57] Adessive plus doubled genitive.

III	a	[+important]	d	[αimportant]
IV	d	[-important]	d	[αimportant]

Accordingly, there is a sharp clear cut between the first three classes (prototypically [+anim;+singular]) (*-na*) and the rest, including all plurals (*-ən*). As with genitives, the dichotomy is canceled in the oblique cases (*-ni*), cf.:

(44)

ABS	SG	PL	
I	jug-na gade	jug-un gade-bi	"nice boy"
II	jug-na xunaʃːe	jug-un jedar[58]	"nice woman"
III	jug-na ɢarg	jug-un ɢarg-əbə	"beautiful wether"
IV	jug-un χaw	jug-un χaj-bə	"beautiful house"

ERG	SG	PL	
I	jug-ni gadeː	jug-ni gade-biʃːe	"nice boy"
II	jug-ni xunaʃːeː	jug-ni jedaːʃːe	"nice woman"
III	jug-ni ɢarg-an	jug-ni ɢarg-əbiʃːe	"beautiful wether"
IV	jug-ni aw-an	jug-ni χaʲ-biʃːe[59]	"beautiful house"

The NPs showing this type of agreement mostly involve so-called relational adjectives, some of which are simply nominal genitives[60]. Especially with loan adjectives which can also be used without any agreement morpheme, this type of classification can indicate definiteness, cf.[61]:

(45) kasib adamij "poor man"
kasib-na adamij "the poor man (of whom I talked)"

The genitive based type of agreement within NPs competes with the original, East Caucasian paradigm of noun classification, which in Tsakhur has undergone considerable changes. This type, already addressed in chapter 4.2.1 appears in two variants: 1) class marker as a prefix, 2) class marker as a suffix. The so-called short adjectives in Tsakhur are marked by suffix CM plus a final *-a* (probably an old auxiliary **-q'a*); they often appear as predicative adjectives or as adverbs, cf.:

(46)

	SG	PL	
I	jug-r-a	jug-b-a	"(is) good"
II	jug-r-a	jug-b-a	
III	jug-b-a	jug-d-a	
IV	jug-d-a	jug-d-a	

[58] *jedar* (itself *jed-ar* "mothers") is the suppletive plural of *xunaʃːe* "woman".
[59] *χaw* "house" has its labial regularly changed to *-j-* before the plural morpheme.
[60] Examples are numerous, cf. *itːuna* "sweet" (*itʷ* "honey"), *boʃuna* "deaf" (*boʃ* "deafness"); in some cases these adjectives are derived participles (*əkːanna* "beloved", *dikaranna* "small", *ʁənanna* "drunken").
[61] The suffix *-na* ~ *-n* also appears with demonstratives to indicate exact reference, cf. *hama-na* "just that one", *ina-na* "just this one" etc.

In the plural, the class marking follows the canonical type of East Caucasian (cf. again (22)). An innovation of Tsakhur is the syncretism of the singular classes I and II, which creates a subclass [+human]. A restricted number of adjectives combines the two procedures mentioned so far, cf.:

(47)

	SG	PL	
I	ts'ele-r-na adamij	ts'ele-b-ən adam-er	"naked man"
II	ts'ele-r-na xunaʃːe	ts'ele-b-ən jedar	"naked woman"
III	ts'ele-b-na balkan	ts'ele-d-ən balkan-ar	"miserable horse"
IV	ts'ele-d-ən ɢelʲ	ts'ele-d-ən ɢelʲ-bə	"naked leg"

Contrary to the genitive based class marker, the CM of the East Caucasian type are preserved in the oblique cases, too (cf. *ts'elemi adameː*, *ts'elemi xunaʃːeː* etc.). Used as prefixes, however, a decisive sound change has taken place, which led to the syncretism of class I and class IV, cf.:

(48)

	SG	PL	
I	Ø-ək'ar-na adamij	w-uk'ar-an adam-er	"ill man"
II	j-ik'ar-na xunaʃːe	w-uk'ar-an jedar	"ill woman"
III	w-uk'ar-na balkan	Ø-ək'ar-an balkan-ar	"ill horse"
IV	Ø-ək'ar-an ɢelʲ	Ø-ək'ar-an ɢelʲ-bə	"ill leg"

Synchronically we thus can observe two different types of classification based on the paradigm of East Caucasian class markers. Whereas the suffixed type probably has some semantic reality, the classification by prefixes is purely formal (at least in the singular). When nomina-lized the adjectives of most SECL undergo quite complex operations of case inflection, cf. the paradigm given for Tsakhur *ts'er-na* "the new one" (here only class I is given for the cases):

(49)

	SG					PL				
		CM_1	CM_2	SA	Case		CM_1	CM_2	SA	Case
ABS	ts'e	-r	-na	-Ø	-Ø	ts'e	-b	-ən	-Ø	-Ø
ERG	ts'e	-r	-un	-ɢʷ	-eː	ts'e	-b	-ən	-biʃː	-e
GEN	ts'e	-r	-un	-ɢʷ	-CM_2	ts'e	-b	-ən	-biʃː	-CM_2
DAT	ts'e	-r	-un	-ɢʷ	-əs	ts'e	-b	-ən	-biʃː	-is

The whole paradigm without further analysis is given in (50):

(50)

SG		I	II	III	IV
ABS		ts'erna	ts'erna	ts'erna	ts'erən
ERG		ts'eruɢʷe	ts'erəɢe	ts'ebəntʃin	ts'edəntʃin
GEN	I	ts'eruɢuna	ts'erəɢəna	ts'ebəntʃina	ts'edəntʃina
	II	ts'eruɢuna	ts'erəɢəna	ts'ebəntʃina	ts'edəntʃina
	III	ts'eruɢuna	ts'erəɢəna	ts'ebəntʃina	ts'edəntʃina
	IV	ts'eruɢun	ts'erəɢən	ts'ebəntʃin	ts'edəntʃin
	PL	ts'eruɢuni	ts'erəɢəni	ts'ebəntʃini	ts'edəntʃini

DAT		ʦ'erunɢus	ʦ'erəɴɢəs	ʦ'ebəntʃis	ʦ'edəntʃis
PL		I	II	III	IV
ABS		ʦ'ebən	ʦ'ebən	ʦ'edən	ʦ'edən
ERG		ʦ'ebənbiʃe	ʦ'ebənbiʃe	ʦ'edəntʃiʃe	ʦ'edəntʃiʃe
GEN	I	ʦ'ebənbiʃda	ʦ'ebənbiʃda	ʦ'edəntʃina	ʦ'edəntʃina
	II	ʦ'ebənbiʃda	ʦ'ebənbiʃda	ʦ'edəntʃina	ʦ'edəntʃina
	III	ʦ'ebənbiʃda	ʦ'ebənbiʃda	ʦ'edəntʃina	ʦ'edəntʃina
	IV	ʦ'ebənbiʃin	ʦ'ebənbiʃin	ʦ'edəntʃin	ʦ'edəntʃin
	PL	ʦ'ebənbiʃdi	ʦ'ebənbiʃdi	ʦ'edəntʃini	ʦ'edəntʃini
DAT		ʦ'ebənbiʃis	ʦ'ebənbiʃis	ʦ'edəntʃis	ʦ'edəntʃis

(50) shows that - as in many other ECL - the nominalization strategy in Tsakhur operates on a decisive dichotomy in the oblique cases, which separates humans from non humans in both numbers. Again, humans are marked for [±mask]. The following morphemes - used as stem augments - appear:

(51)	SG	PL
[+mask;+hum]	-ɢʷ-	-ʃ-
[-mask;+hum]	-ɢ-	-ʃ-
[-hum]	-ʧi-	-ʧi-

This type of nominalization also works for pronouns, numerals and so on and represents one of the basic strategies to classify referents in oblique functions. It can be shown that the stem augments probably represent older ergative morphemes[62] which led to a secondary type of classification competing with that one used with absolutives. The opposition SA[+mask] *-ɢʷ-* vs. SA[+fem] *-ɢ-* is obviously based on a strategy to mark a "male connotation" through labialization (whereas the "female" semantics is left unmarked). Perhaps this labial element is related to the (old) class marker PL **-w-* ([+mask;+hum]), though its position would be quite unexpected. One assumption is that this morphemic cluster results from a time when the proto-language (PEC ?) did not discriminate the functional cases ergative, genitive, and dative, instead using an obliquus for all of them. Hence we would have had a pair ABS *NOUN-∅ vs. OBL *NOUN[+mask;+hum]-*ɑ(ə)-w*. However, the fact that the feminine (CL II) is left unmarked (instead of a hypothetical form **-ɢʲ* - < **ɑ(ə)-j*) seems to question this assumption[63].

[62] An assumption that is based on the hypothesis that once the East Caucasian proto-language had a simple category "obliquus", which stood in binary opposition against the absolutive (*casus rectus*), and which was subcategorized lateron (cf. Topuria 1995 for details).

[63] It should be noted that the opposition "labial" > [+mask] vs. "non labial" > [+fem] can also be traced in the lexicon of Proto-Lezgian. Thus we can reconstruct a form **ɨʷə-* "man" (e.g. Aghul *xuj* ~ *ʃuj* ~ *ʃʷuj*, Kryts *firi*, Budukh *furi*, Udi *iɑ(u)*, perhaps Tsakhur *jer* "man") which stands against **ɨə-(-n/d-)* "woman" (> Tsakhur *xuna-ʃːe* "woman", *xər-un* (< **xəd-un*) "little woman").

4.4 Pronouns

4.4.1 Personal Pronouns

The paradigm of personal pronouns in Tsakhur should be treated in the light of Proto-Lezgian. In general, the system of indexing speech act participants (SAP) in SECL is based on the distinction between <SPEAKER> and <HEARER>, both singular and plural. Historically, the proto-language knew the discrimination of "inclusive" ("I/we plus you") vs. "exclusive" ("I/we plus them (without you)"). This system has been structurally preserved in Tabasaran, Aghul, Khinalug, Archi, and (partly) in Southern Samur, though Archi (and perhaps Khinalug) have developed formally new inventories[64]. As for case marking, some SECL show the (functionally expected) syncretism of absolutive and ergative[65]. It must be said, however, that the reconstruction of the inflectional paradigm of personal pronouns in Proto-Lezgian hints at a quite systematic differentiation between absolutive (marked **-∅* or **-n*) and ergative (marked by the "pronominal ergative" **-a*)[66]. In this respect the Tsakhur dialect behaves quite innovating: It has given up both the inclusive/exclusive and the case dichotomy. The paradigm is give in (52):

(52)	ABS	ERG	GEN
1	zə	zə	jiz-da / jiz-ən / jiz-di
2	ʁu	ʁu	jiʁ-na / jiʁ-ən / jiʁ-ni
4	ʃi ~ ʃːi	ʃi ~ ʃːi	jiʃ-da / jiʃ-in / jiʃ-di
5	ʃu	ʃu	wuʃ-da / wuʃ-ən / wuʃ-di

Table (53) contrasts this system with the paradigm proposed for Proto-Lezgian:

(53)	Tsakhur (Ts.)		Proto-Lezgian	
	ABS	ERG	ABS	ERG
1	zə	zə	*zʷə-	*zʷə-a
2	ʁu	ʁu	*ʁʷə-	*ʁʷə-a
4i	ʃi ~ ʃːi	ʃi ~ ʃːi	*xːə-	*xːə-a
4e	---	---	*ʒʲə-	*ʒʲə-a
5	ʃu	ʃu	*ʒʷə-	*ʒʷə-a

The dialect of Gelmets has reestablished the ABS/ERG-dichotomy (perhaps under influence from Rutul), cf.[67]:

[64] For the system of personal pronouns in the SECL and in Proto-Lezgian see Schulze 1997b and the bibliography given there.

[65] This "expectation" is based on the well-known Silverstein Hierarchy and is well attested in many ergative languages. The basic claim is that SAP are prototypically agentive and that in the cognitive representation of States of Affair the "view point" is naturally based on the (agentive) SAP if present.

[66] Cf. Schulze 1997a for details. Within SECL, a dichotomy ABS/ERG (sometimes restricted to certain SAP) is found in Lezgi, Tabasaran (only in verbal inflection), Aghul (only dialect of Richa), Rutul (partly), Archi (only SAP(1)), and Khinalug (only singular).

[67] See Ibragimov 1990:193-194, ʒeiranišvili 1984:558.

(54) Tsakhur (Gelmets):

	ABS	ERG	GEN
1	zə	za-sːa	zəni
2	ʁu	wa-sːe[68]	jəˀʁˀna / -ən
4	ʃːi	ʃːa-sːa	jiʃna / -ən
5	ʃu	ʃʷa-sːæ[69]	juʃna / -un

The forms of the ergative are isomorphic with those e.g. in Mukhad (Rutul). The (innovated) suffix *-sːɪ ~ -sːe* follows the old pronominal ergative marker *-a*, the distribution of which still reflects the Proto-Lezgian conditions. As compared to Tsakhur (dialect of Tsakhur) the Gelmets paradigm seems to be older (though the ergative morpheme itself is surely an innovation). Tsakhur has reduced the whole system, whereas Gelmets still strengthened the ergative organization of the paradigm. For Proto-Tsakhur we can assume the following system of personal pronouns:

(55) Proto-Caxur:

	ABS	ERG
1	*zə	*zə-a
2	*ʁʷə	*ʁʷə-a
4i/e	*xːi	*xːi-a
5	*ʃu	*ʃu-a

4.4.2 Demonstrative Pronouns

Just like the system of personal pronouns, the demonstrates should be explained with the help of Proto-Lezgian. The prototypical paradigm of demonstratives in the SECL represents a system of indexing trajectors in landmarks established from the point of view of the speaker. Historically, they were not sensitive for any SAP orientation. As many other ECL some SECL subcategorize the (horizontal) distal relation vertically, referring to entities far away from the speaker, but higher or lower as himself. This system is not preserved in all SECL, its basic paradigm for Proto-Lezgian is given in (56)[70]:

(56)		
	Proximal	*mV-
	Distal (horizontal)	*dV-
	Distal (higher)	*ɢu- / *ʁu-
	Distal (lower)	*gu-

[68] Surely developed from **ʁʷa-sːe*.
[69] Kibrik/Kodzasov 1990:223 mention a 2.PL ERG only for Mikik (*ʃo-sːe*). ǯeiranišvili 1984:559 has *ʃe* (< **ʃi-æ*). The form *ʃʷæsːæ* was confirmed by my informants.
[70] Cf. Schulze 1997b for details.

Tsakhur uses the following deictic pronouns which, however, no longer reflect a vertical orientation[71]:

(57)				
	1	ina (I-III) / in (IV)	SAP(1)	"this here"
	1'	hajna (I-III) / hajin (IV)		"this exactly here"
	2	mana (I-III) / man (IV)	SAP(2)	"that"
	2'	hamana (I-III) / haman (IV)		"exactly that"
	3	ʃena (I-III) / ʃen (IV)	¬SAP	"that over there"
	3'	hoʃina (I-III) / hoʃin (IV)		"exactly that over there"

The basic structure reveals a tripartite system, which historically was not dependent of the location of speech act participants, but reflected the speaker's subjective interpretation of a referent's location, cf.[72]:

(58)		
	Proximal	i-
	Medial	ma-
	Distal	ʃe-[73]

ma-n- / ʃe-n- also encode the opposition [+visible] (*man-*) vs. [-visible] (*ʃen-*). The inflection of the nominalized demonstrative pronouns is basically that of the adjectives (cf. 4.3) The table in (59) lists the absolutive, ergative, and genitive forms of the pronoun *ma-n-* (the other pronouns are inflected analogically):

(59) ma-n- [2, medial, [+visible]]:

	ABS	ERG	GEN
I	ma-na	man-ʁʷ-e [74]	man-ʁu-na / -n
II	ma-na	man-ʁ-e	man-ʁə-na / -n
III	ma-na	man-ʧə-ən	man-ʧə-na / -n
IV	ma-n	man-ʧə-ən	man-ʧə-na / -n

4.4.3 Reflexives[75]

Syntactic reflexivity in quite marginal in the SECL, as generally expected for patient oriented systems of ergativity. Still, every SECL has a specific reflexive pronoun that is basically used to emphasize the role of the agent (much rarer that of the patient or addressee). Lezgian lan-

[71] There are many free variants of these demonstratives (e.g. *himan-* instead of *haman-*) which are hard to systematize.
[72] To all pronouns secondary emphatic particles can be added.
[73] Tsakhur *ʃe-* forms an important isoglosse with Udi *ʃe-* (distal, only nominalized).
[74] The SA *-ɢʷ* and *-ɢ-* often become *-ʁʷ-* and *-ʁ-* after nasals.
[75] Comparative research on reflexives in the SECL is still needed. Some general (and preliminary) informations on the single language paradigms, on the degree of their functionality, and on their Proto-Lezgian background is given by Schulze 1997b.

guages that tend towards a strong accusativization of their operating system often use this pronoun as a subject based, "true" reflexive[76].

In some languages the reflexive pronoun is subcategorized either according to the question whether it refers to SAP or not, or with respect to semantic criteria ([±human] or [±mask]). In the absolutive class differentiation is common, though not always functionally preserved. A good example is the system of Rutul (Mukhad)[77], which is based on the pronoun *-ʤ̑ʷə̑- (absolutive)[78]:

(60)

	I	II	III	IV	PL
ABS	wuʤ	riʤ	wiʤ	jiʤ	ʤʷær
ERG	ʤu-da	ʤi-dæ	ʤi-dæ	ʤi-dæ	ʤʷær-ʃ-æ
GEN	ʤu-du	ʤi-də	ʤi-də	ʤi-də	ʤʷær-ʃ-də
DAT	ʤu-s	ʤi-s	ʤi-s	ʤi-s	ʤʷær-ʃ-is

In Tsakhur the paradigm is somewhat more complex. First, the class differentiation shows a very exceptional distribution of Class I+III vs. II+IV (absolutive singular only)[79]; second, the sexus differentiation of human referents is kept in the ergative, and third, there is a strong dichotomy to be observed between human and non human referents in the oblique cases (*ʤ̑ʷə̑ʲ- vs. *ʧi-). As in many other SECL the inflectional paradigm is that of nouns. Also note that the OBL of non human referents (ʧi-) is the same as the SA [-hum] (-ʧi-), cf.

(61)

SG	I	II	III	IV
ABS	wuʤ[80]	jiʤ	wuʤ	jiʤ
ERG	wuʤ-eː	jiʤ-eː	ʧi-n	ʧi-n
GEN	ʤu-nan	ʤe-nan	ʧi-nan	ʧi-nan
	ʤu-ni	ʤe-ni	ʧi-ni	ʧi-ni
DAT	ʤu-s	ʤe-s	ʧi-s	ʧi-s

PL	I/II	PL III/IV
ABS	ʤo[81]	jiʤ-bə
ERG	ʤo	ʧi-n
GEN	ʤo-nan	ʧi-nan

[76] On this cf. chapter 5.3.
[77] Cf. Ibragimov 1978:79f.
[78] Borch-Khnov has neutralized the class differentiation on the basis of the highly unmarked class III (*wiʤ*).
[79] This distribution is based on phonetic reasons only: CL III < **b-uʤ*, CL IV < **d^L-iʤ* < * *d-iʤ*). But note that the reflexive pronoun is the only instance in which the old PL distinction of CL I and CL II markers is still preserved. The vocalism in the oblique cases reflects this opposition: The ergative form had its accent on the root whereas the other case forms were accentuated on the suffix. This let to the loss of the first vowel and to the reduction of the second one which then was colored according to the phonetics of the CM (**wə-ʤə-nán* > **u-ʤə-nán* > *ʤu-nán* etc.).
[80] Gelmets has ABS *иʤ*, ERG *иʤе*, perhaps composed out of **иu-* + **-ʤə-*. Else, the paradigm equals that of the dialect of Tsakhur.
[81] *ʤo* is a reflex of the "simple" stem *-ʤ̑ʷə̑- joint by the plural morpheme *-a* (> **ʤ̑ʷa* > *ʤo*). The hypothetical plural ergative **ʤ̑ʷə̑ʲ-a-e* (> **ʤ̑ʷa-e* > **ʤo-e*) has been substituted by the absolutive form out of phonetic reasons.

	ʤo-ni	ʧi-ni
DAT	ʤo-s	ʧi-s

4.4.4 Interrogative Pronouns

As most other Lezgian languages, Tsakhur has to different paradigms of referential interrogatives, encoding the opposition [±human]. In the plural only the absolutive pronouns are marked, which is another argument for the basically patient oriented ergativity in Tsakhur, see below. The paradigms for "who" and "what" are given in (62):

(62)

	SG		PL	
	"who"	"what"	"who"	"what"
ABS	haʃːu-ne	hiʤoː-ne	haʃ-bə-ne	hiʤo-bə-ne
ERG	ʃaw-a-ne[82]	niʃ-e-ne	ʃaw-a-ne	niʃ-e-ne
GEN	ʃaw-na-ne	niʃ-ina-ne	ʃaw-na-ne	niʃ-ina-ne
	ʃaw-ən-ne	niʃ-in-ne	ʃaw-ən-ne	niʃ-in-ne
	ʃaw-ni-ne	niʃ-ini-ne	ʃaw-ni-ne	niʃ-ini-ne
DAT	ʃaw-us-ne	niʃ-is-ne	ʃaw-us-ne	niʃ-is-ne

The absolutive forms are marked by a deictic element *ha-*, which has become *hi-* in the non-human question word out of analogy (OBL *niʃ-*). The final morpheme *-ne* is an innovation of Tsakhur. With the help of this element the speaker directly questions a SAP(2), whereas *-je* asks for a ¬SAP (via SAP(2)), cf.:

(63) haʃːu-ne maː
who.ABS-Q there
"Who is there?"

hoːra ∅-aʦ'a-xe haʃːu-je akːa-s
I.go.IMP I-know.PRES-COND who-Q dooryard-AD.ESS
"Go (to) know who is (there) in the dooryard!"[83]

The element *ne-* also forms the basis for the expression of "which (one)". The interrogative *nen-* is composed out of the original question particle *ne-*, to which the attributive (definite) class markers (*-na /-n*) are added. It is inflected like nominalized adjectives, cf.:

[82] The old pronominal *a*-ergative is preserved with (potential) human referents. The use of the nominal ERG[+hum] (*-e*) for "what" ([-hum]) is probably based on the fact that asking for a potential agentive non human being still refers to its possible function to control a transitive action.
[83] Both particles also function as indicators of yes/no-questions (cf. *χaːj-e dakʲ* (house.LOC-Q father.ABS) "Is the father at home?"). They are then cliticized to the word questioned, and differ in this respect e.g. from the Udi question particle *-a*, which is restricted to third persons.

(64) "Which (one)"

	SG I	II	III	IV
ABS	nena-ne	nena-ne	nena-ne	nen
ERG	nen-ɢ^{w}-eː-ne	nen-ɢ-eː-ne	nen-ʧi-ne	nen-ʧi-ne
GEN	nen-ɢu-na-ne	nen-ɢ-əna-ne	nen-ʧi-na-ne	nen-ʧi-na-ne
	nen-ɢu-n-ne	nen-ɢ-ən-ne	nen-ʧi-n-ne	nen-ʧi-n-ne
	nen-ɢu-ni-ne	nen-ɢ-ni-ne	nen-ʧi-ni-ne	nen-ʧi-ni-ne
DAT	nen-ɢu-s-ne	nen-ɢ-əs-ne	nen-ʧi-s-ne	nen-ʧi-s-ne

	PL I + II	III + IV
ABS	nen-bə-ne	nen-bə-ne
ERG	nen-bi-ʃ-e-ne	nen-ʧi-ʃ-e-ne / nen-ʧi-n-ne
GEN	nen-bi-ʃ-da-ne	nen-ʧi-na-ne
	nen-bi-ʃ-in-ne	nen-ʧi-n-ne
	nen-bi-ʃ-di-ne	nen-ʧi-ni-ne
DAT	nen-bi-ʃ-is-ne	nen-ʧi-s-ne

4.5 Numerals

In this section first the cardinals for 1 - 10 and 20 are given together with the reconstructed forms of Proto-Lezgian (dialectal variants are mostly neglected):

"One":
PL *sa-: Lezgi *sad*, Tabasaran *sa-w*, Aghul *sːar / sad*, Rutul *sa-*, **Tsakhur *sa-***, Archi *os* (ATTR *sːej-*), Kryz *sa-*, Budukh *s-*, Khinalug *sa-*, Udi *sa*.

"Two":
PL *$q^{ʕw}$'ə-: Lezgi q^{w}'æd, Tabasaran $q^{ʕ}$'u-, Aghul $q^{ʕ}$'ur, Rutul $q^{ʕw}$'a-, **Tsakhur *$q^{ʕ}$'o-***, Archi $q^{ʕw}$'e-, Kryz q^{w}'a-, Budukh *q'a-*, Khinalug *k'u-*, Udi *p'a*^{ʕ}.

"Three":
PL *łəbə-: Lezgi *pːud*, Tabasaran *ʃibbu-*, Aghul *fːubuːr* (Burkhikan), *xibud* (Richa), Rutul *xʲibə-*, **Tsakhur *xeb-***, Archi. *łeb-*, Kryz *ʃibə-*, Budukh *ʃubu-*, Khinalug *p^{h}f^{w}a-*, Udi *χib*.

"Four":
PL *j/bəq^{w}'ə-: Lezgi *q'ud*, Tabasaran *jaqː'u-*, Aghul *jaq'ud/r*, Rutul *juqu-*, **Tsakhur *joqu-***, Archi *ebq'a-*, Kryz *juq'u-*, Budukh *juq'u-*, Khinalug *p^{h}xu-*, Udi *bip'*.

"Five":
PL *f^{w}ə-: Lezgi *wad*, Tabasaran *xu-*, Aghul *jif^{w}ur / ʕafud*, Rutul *xu-*, **Tsakhur *xo-***, Archi *łːo-*, Kryz *fə-*, Budukh *fu-*, Khinalug *pxu-*, Udi *qo*.

"Six":
PL *(jə)rəɫ^{w}ə-: Lezgi *rəgid*, Tabasaran *jirxu-*, Aghul *jerxid*, Rutul *rixe-*, **Tsakhur *jixe-***, Archi *dil-*, Kryz *rəxə-*, Budukh *rəxə-*, Khinalug *zæk^{h}-*, Udi *u*^{ʕ}*q*.

"Seven":
PL *w/jərɣ-: Lezgi *erid*, Tabasaran *urʁu-*, Aghul *yerer/d*, Rutul *jiwe-*, **Tsakhur *jiye-***, Archi *witʲ-*, Kryz *jiyə-*, Budukh *jiji-*, Khinalug *jik'*, Udi *wu*^{ʕ}*ʁ*.

"Eight":
PL *mərɣ-: Lezgi *miʒid*, Tabasaran *miʒi-/miʒːu-*, Aghul *mujar/d*, Rutul *məje-*, **Tsakhur *moli-***, Archi *metʲe-*, Kryz *miye-*, Budukh *məjə-*, Khinalug *ʔik'*, Udi *mu*^{ʕ}*ʁ*.

"Nine":
PL *y/wətʃ^{w}'-: Lezgi *tʃ'yd*, Tabasaran *hurtʃ^{w}'u-*, Aghul *jertʃ^{w}'u-*, Rutul *hutʃ'u-*, **Tsakhur *ju^{ʕ}tʃ'u-***, Archi *utʃ'-*, Kryz *jitʃ'i-*, Budukh *witʃ'i-*, Khinalug *joz*, Udi *wuj*.

"Ten":
PL *y/wəts'-: Lezgi *ts'yd*, Tabasaran *jits'u-*, Aghul *jic'ur/d*, Rutul *jits'e-*, **Tsakhur *jitsə'-***, Archi *wits'-*, Kryz *jits'ə-*, Budukh *jəts'ə-*, Khinalug *jæʔiz*, Udi *wits'*.

"Twenty":
PL *qːa-/*ɢa-: Lezgi *qːad*, Tabasaran *qːa-/ɢa-*, Aghul *qːur/d*, Rutul *ɢaː-*, **Tsakhur *ɢa-***, Archi *q'a-*, Kryz *ɢa-*, Budukh *qa-*, Khinalug *q'an-*, Udi *q'a*.

The numerals 21 and higher are formed on the basis of *ɢa-* (20), *xebts'al^{j}* (30), *joq'ts'al^{j}* (40), *xots'al^{j}* (50), *jəxts'al^{j}* (60), *jiyts'al^{j}* (70), *molʲts'al^{j}* (80), and *ju^{ʕ}tʃ'a^{ʕ}l^{j}* (< *ju^{ʕ}tʃ'-ts'al^{j}) (90). For 11-19 the following forms are used: *jits'əssːa-* (11), *jits'əq^{ʕw}'a-* (12), *jits'əxeba-* (13), *jits'aːq^{w}'a-* (14), *juts'uxːa-* (15), *jits'ijxa-* (16), *jits'ijya-* (17), *jits'əmole-* (18), *jits'ijtʃ'^{w}a^{ʕ}-* (19). The basic counting system is TEN+X for the tens, and X-TEN-SUPER.ESS-X (e.g. "three on ten [and] two" = 32) for the twenties and higher[84].

When used attributively all numerals except *wa^{ʕ}ʃ* "hundred" add the morpheme *-le* to the class marked stem, which is assimilated if possible. Thus we have *q^{ʕ}'o^{ʕ}-b-le zer* "two cows", *xej-ib-le zer* "three cows" etc. In the oblique cases, class markers and *-le* are replaced by the OBL morpheme *-ni* (*q^{ʕ}'o^{ʕ}-ni / xeb-ni zer*). As it can be seen the numerals do not invoke a nominal plural which is an economic procedure common to most SECL (except for Udi which sometimes uses plurals with animate nouns).

Ordinal numbers are indicated with the help of the suffix *-esda* (I-III) ~ *-esən* (IV) which - as in many ECL - is derived from the verb *e-hes* "to say". The numeral itself is inflected by class markers. Thus we have:

(64)	I	sa-r-esda	"The first one"
	II	sa-r-esda	
	III	sa-p'-esda < *sa-b-hesda	
	IV	sa-t'-esən < *sa-d-hesda	

In many instances, the Tsakhur ordinals are replaced by their Russian or Azeri counterparts.

[84] Ibragimov 1990:94 assumes that the element *-lʲ* in *-ts'alʲ* results from a "neutralized CM", which is neither structurally nor morphologically understandable. The prefix (?) *ji-* is dropped in the decades.

4.6 Adpositions

As all other SECL languages Tsakhur knows a quite complex system of (mostly local) postpositions. In many cases they can be regarded as grammaticalized nouns, often additionally marked by (local) case suffixes. (65) lists some of these postpositions which have their nominal head either in the genitive or in a (appropriate) local case:

(65)	"under"	awu-r (I/II), awu-b (III), awu- (IV)
	"on"	o
	"onto"	oːqa
	"from above"	oːnʧe
	"between"	æræ
	"after"	qijʁa
	"aside"	suralʲ
	"in"	ar (I/II), ab (III), ad (IV)
	"till"	ʁana

The class marked postpositions are triggered by a noun in the absolutive ("class attraction"), cf.

(66) dakʲ χaj-ən a-r wo-r-or
father(I).ABS house-GEN in-I AUX-I-PRES.CONF.I
"Father is at home."

kitab sanduχ-ən a-b wo-b-ob
book(III).ABS box-GEN in-III AUX-III-PRES.CONF.III
"The book is in the box."

baluʁ xʲan-ən a-d wo-d-od
fish(IV).ABS water-GEN in-IV AUX-IV-PRES.CONF.IV
"The fish is in the water."

4.7 Adverbs

Most of the adverbs in Tsakhur are derived from either adjectives or nouns marked by a locative (such as *ɢij-na* "today", *ɢij-qa* "tomorrow", *əxa-lʲ* "in the evening", and many local adverbs). However, there are some (tough very few) non derived adverbs, e.g. *ɢʲej-* "much", *sana-* "together", *ʦ'ele-* "naked", *dʲoles* "near", *k'ane* "near", *qijha* "in the future". Adverbs derived from adjectives are marked by a suffix CM (*-r, -r, -b, -d*) instead of the attributive suffix *-na / -Vn* (cf. *ək'-r-a* (I,II), *ək'-b-a* (III), *ək'-d-a* (IV) "quickly" (vs. *ək'-na ~ ək'-ən* "quick"). The class marked adverbs always agree with the syntactic function "subject", i.e. with the absolutive marked noun phrase, cf.

(67) deχ ək'-ra maktab-eːqa qarə
son(I).ABS quickly-I school-IN.ALL I.come.PAST
"The son came quickly to school."

balkan ɢʲej-ba w-uk'ar-o-b
horse(III).ABS very-III III-ill-AUX-III
"The horse is very ill"

zə hajna kaʁəz ək'-ba ojk'an-nij
I.ABS this letter(III).ABS quickly-III III.write-PAST.1.SG
"I have quickly written this letter."

zaz ɢʲej-da d-əkːanan-bi jizən ʧoʒ-er-əj jiʧi-bi-j [85]
I.DAT much-I/II.PL I/II.PL-love.PRES-PL my brother-PL-and sister-PL-and
"I love my brothers and sisters very much"

4.8 Verbs

In an optimized account the verbal paradigm in the SECL is characterized by the following, often interacting features: A verbal stem which in many cases consists only of a quite minimal phonetic body (V, C, VC, CV) is marked for tense, aspect, aktionsart, mode, localization, pluralization, and agreement. Historically, all these categories except aspect were indicated by agglutinating prefixes or suffixes, aspect, however, was encoded either by root ablaut or by a specific set of infixes[86]. Personal inflection is quite marginal, it can be found fully paradigmatic in Tabasaran and Udi, partly in Tsakhur, and - differently motivated - in many paradigms of the imperative mood.

4.8.1 Stem Formation

In Tsakhur, as well as in the other SECL, there is only a small number of underived verbal stems. Mostly, they are marked by preverbs which originally denoted locational aspects, but which are often lexicalized today (see below). There is clear evidence for what is called "thematic stem vowel" in Tsakhur, a phenomenon that is highly characteristic for Eastern Samur. Secondary stems are formed with the help of auxiliaries, which sometimes incorporate nouns or adjectives. In Tsakhur, *haʔas* "to make" and *ixes* "to be(come)" are used in this respect.

4.8.2 Class Markers

Class markers play an important role in the agreement system of all SECL that differentiate noun classes morphologically. Normally, they are conditioned by the absolutive NP in a phrase (see chapter 5.2 for details). In Tsakhur class markers interact in a very complicated way with

[85] This is an Gelmets example. The form *d-ekːanan-bi* had been used instead of *b-əkːanan-bi*, which would be the expected verb form (CM *b-* for CL I/II.PL).
[86] A general overview on morphological aspect marking techniques in the SECL is given e.g. by Schulze-Fürhoff 1994c.

other verbal morphemes or with the verbal stem so that it is difficult to give a comprehensive account of this feature here. In fact, it is the position of the class markers which is decisive for its final representation, but we have to bear in mind that the resulting morphophonemic processes already started when the original shape of the class markers (cf. (24)) still was preserved.

Class markers can appear either as prefixes, infixes, or suffixes. In Tsakhur the position between preverbs, if present, aspect markers, and stem initials. (67) gives a simple example for *alʲa-CM-t'-as* (INF) "to carry":

(68)		SG	PL
	I	alʲa-r-t'-as	alʲa-b-t'-as
	II	alʲa-r-t'-as	alʲa-b-t'-as
	III	alʲa-b-t'-as	alʲa-Ø-t'-as
	IV	alʲa-Ø-t'-as	alʲa-Ø-t'-as

The underlying paradigm is given in (69):

(69)		SG	PL
	I	*ala-w-t'-as	*ala-b-t'-as
	II	*ala-r-t'-as	*ala-b-t'-as
	III	*ala-b-t'-as	*ala-d-t'-as
	IV	*ala-d-t'-as	*ala-d-t'-as

When the class marker is the first element in a verbal chain, then there is a different type of class distribution, showing the merger of class I and IV, cf. the verb CM-*əx-es* (INF) "to become":

(70)		SG	PL
	I	Ø-ix-es	w-ux-es
	II	j-ix-es	w-ux-es
	III	w-ux-es	Ø-ix-es
	IV	Ø-ix-es	Ø-ix-es

The phonetic processes that result from an interaction of class markers and aspect morphemes leads to highly lexicalized types of ablaut, cf. as an example *k'e*-CM-*(j)x-as* (INF) "to fall" (here *k'e-je*-CM-*x-a* (present tense durative), only the singular forms are given, as the plurals can easily be derived from them):

(71)				
	I	kʲ'-e-x-a	<	*k'í-Ø-əj-x-a
	II	kʲ'-eː-x-a	<	*k'í-j-əj-x-a
	III	kʲ'-oː-x-a	<	*k'í-w-əj-x-a
	IV	kʲ'-e-x-a	<	*k'í-Ø-əj-x-a

The class marking types given in (69) and (71) sometimes appear as free variants within the same paradigm, which hints at still relatively unstable conditions, cf. for *gʲoqas* "to suck" (*gʲoqu* "sucked"):

(72)				
	I	gʲ-o-q-u	~	gʲ-o-r-q-u
	II	gʲ-eː-q-u	~	gʲ-o-r-q-u
	III	gʲ-oː-q-u	~	gʲ-o-b-q-u
	IV	gʲ-o-q-u	~	gʲ-o-d-q-u

4.8.3 Preverbs

The system of preverbs is much more productive in the Nothern SECL than e.g. in Southern Samur or in Udi. Tabasaran has strongly grammaticalized this pattern by using it for aspectual differentiation[87]. The material of preverbs is somewhat related to the local cases suffixes, though there is not always a complete identity. In Tsakhur we meet for instance the following preverbs:

(73)		
	VV-	"away from the speaker"
	VlV-	"up"
	Vt'V-	unsure meaning
	Vts'V-	unsure meaning
	atʃ'i-	unsure meaning
	sV-	"towards something" (= adessive)
	ɢV-	"action away from the speaker"
	gV-	"down"
	k'V-	"down into"
	qV-	"action separating something from something"
	hV-	"away from something"

The functions of these preverbs indicated by the glosses cannot be taken for granted. In many cases - as said above - the preverbs are highly desemantisized, forming a lexical unity with their verbal stem. Moreover, many verbs show preverbial clusters, combining different preverbs or using them with now obsolete locational morphemes. (68) gives the paradigm for -*k'al/n-as* "to jump", which is still sensitive for locational preverbs (only class I is given here):

(74)		
	o-k'alas	"to jump"
	ɢo-k'alas	"to jump over something"
	kʲ'o-k'anas	"to jump down into the deep"
	ilʲo-k'alas	"to jump down"
	ilʲqo-k'olas	"to jump up repeatedly"

[87] In a way that is fairly common in Slavic languages. Forms with preverb denote a perfective aspect, whereas the bare forms are used to indicate imperfectiveness.

4.8.4 Tense and Aspect

As in most other SECL there is a strong interdependency to be observed between tenses forms and aspect markers in Tsakhur. In fact, it is often hard to discriminate both categories morphologically. The structure of TAM paradigms in the SECL is often based on a tripartite system, which can be described as follows:

(75) INFINITIVE - PRESENT GERUND - PAST GERUND

These three basic forms are followed by secondary morphemes the status of which can hardly be described in the context of PL[88]. Present and past gerund represent the prototype of the aspectual system in SEC ([±durative]). As expected, in Tsakhur the durative is the morphologically marked category. It can be indicated either by infixing *-i-*, *-a-*, or *-r-*, by ablaut (mostly *-i-*, *-a-* > *-e-*; *-a-* > *-i-* (especially for the vowels of preverbs)), by reduplication or by lexical suppletion. The distribution of durative and non-durative forms is conditioned by the semantics of the actual tense form, however, in many cases either the durative or the non-durative stem is lexicalized, so that it is difficult to tell whether a verb in question still operates on the aspectual category or not without knowing the whole paradigm.

The temporal category is of secondary origin in nearly all SECL. It is based on the (old) aspect system, which then is enlarged by different tense morphemes often stemming from case forms, adjective markers and so on. The resulting TAM complexes can play the role of real temporal forms, as participles, or as converbs. The interaction of tense, aspect, and mood leads to often very complicated or at least very sophisticated paradigms, for which Tsakhur gives a good example.

The two basic gerunds are marked by ablaut of the final stem ("thematic") vowel, if present. If the verbal root is augmented by a suffix *-an, -ar*, or *-al*, then the prefinal vowel is changed accordingly. The following system of ablaut operates:

(76)

PRESENT GERUND	PAST GERUND
-a	-u
-a	-ə
-e	-a
-e	-u
-an, -ar, -al	-un, -ur, -ul
-an, -ar, -al	-n, -r, -l (< *-ən, *-ər, *-əl)

The distribution of the ablaut variants with thematic vowels is depending rather on the lexical entry in question than on morphological or phonological reasons (at least on a synchronic level)[89]. The decision for whether a given verb form is a present or past gerund is supported by

[88] As far as I know there is no comparative treatment of the tense systems in the SECL up to now (but see Alekseev 1985, who informs about this topic though rather briefly). A first more general attempt has been made by Xajdakov 1975, though his data are very limited.

[89] Ablaut variation within thematic vowels is one of the basic features of tense formation in the Eastern and Western Samur languages, cf. Moor 1985:69-82 for Lezgi, Ibragimov 1978:102 for Rutul. Aghul and Tabasaran show that the system was originally based on a simple ablaut PRES *-a*, PAST *-u* (cf. Aghul *aq'-a* "making", *aq'-u-ni* (AOR) "made"). The vowel alternations (*-a / -u / -i* etc.) within one tense form are due to secondary processes of assimilation, though in many cases the phonetic motivation is very much obscured. The two markers of the durative aspect *-a* and *-e* may perhaps reflect an old dichotomy [±transitive].

the internal structure of the root: Present gerunds often reflect the (older) morphology of duratives, the morphemes of which are lacking in the past gerund. The fact that the old aspects markers (**-j-*, **-r/l-*) were prefixes led to complexe morphophonological processes due to the interaction of aspect markers, stem initials, class markers, preverbs, and stress. The structure of the Tsakhur verb *ixes* "to be(come)" may clarify this:

(77)

Present Gerund:	CM-íi:-x-e	<	*CM-jə́-əx-e
Past Gerund:	CM-ə́x-a	<	*CM-Ø-ə́x-a
Future/Infinitive:	CM-ix-és	<	*CM-jə-əx-és

The stem initial shwa (*-ə-*) is than effected by the phonetics of a preceding class marker or a preverb, cf. for class markers:

(78)

	PRES	PAST	FUTURE
I	Ø-é:xe	Ø-ə́xa	Ø-ixés
II	j-é:xe	j-íxa	j-ixés
III	w-é:xe	w-úxa	w-ixés
IV	Ø-é:xe	Ø-ə́xa	Ø-ixés

(79) gives an example for a structure "preverb + CM + aspect + verb" (*gʲatʒ'és* "to fill)[90]:

(79)

PRES:	I	gʲájtʒ'e	<	*gʲə́ₐ-Ø-əj-tʒ'e
	II	gʲájejtʒ'e	<	*gʲə́ₐ-j-əj-tʒ'e
	III	gʲáwejtʒ'e	<	*gʲə́ₐ-w-əj-tʒ'e
	IV	gʲájtʒ'e	<	*gʲə́ₐ-Ø-əj-tʒ'e
PAST:	I	gʲátʒ'ə	<	*gʲə́ₐ-Ø-tʒ'ə
	II	gʲájtʒ'ə	<	*gʲə́ₐ-j-tʒ'ə
	III	gʲáwtʒ'e	<	*gʲə́ₐ-w-tʒ'ə
	IV	gʲátʒ'ə	<	*gʲə́ₐ-Ø-tʒ'ə
FUTURE:	I	gʲatʒ'és	<	*gʲəₐ-Ø-j-tʒ'és
	II	gʲajetʒ'és	<	*gʲəₐ-jə-j-tʒ'és
	III	gʲawatʒ'és	<	*gʲəₐ-wə-j-tʒ'és
	IV	gʲatʒ'és	<	*gʲəₐ-Ø-j-tʒ'és

[90] With the verb *gʲatʒ'és* the aspect marker **-jə-* shows metathesis (**-əj-*).

[91] In the future tense the aspect marker **-əj-* (< **-jə-*) has been reduced to *-j-* in prestressed position. It has been dropped then before *-tʒ'*. CL III *gʲawatʒ'és* shows further vowel harmony (under stress, as in the imperative, the form is once more reduced, cf. *gʲa:tʒ'e* "fill!" (< **gʲə́ₐ-wə-tʒ'e*)).

The above mentioned processes are further complicated if other preverbial elements as negation particles appear, or if the class marker enters the root as an infix, cf. for *gʲiwaχanás* "to hang up", which shows an aspect-like prefix *-*wə*- with an obscure history:

(80)	PRES:	I	gʲiwajχan	<	*gʲɔ́-wə-∅-χan
		II	gʲiwe:χan	<	*gʲɔ́-wə-j-əj-χan
		III	gʲiwa:χan	<	*gʲɔ́-wə-w-əj-χan
		IV	gʲiwajχan	<	*gʲɔ́-wə-∅-χan
	PAST:	I	gʲiwarχən	<	*gʲɔ́-wə-r-χən[92]
		II	gʲiwarχən	<	*ʲgʲɔ́-wə-r-χən
		III	gʲiwabχən	<	*gʲɔ́-wə-b-χən
		IV	gʲiwadχən	<	*gʲɔ́-wə-d-χən
	FUTURE:	I	gʲiwaχanás	<	*gʲə-wə-∅-χanás
		II	gʲiwe:χanás	<	*gʲə-wə-j-χanás
		III	gʲiwa:χanás	<	*gʲə-wə-w-χanás
		IV	gʲiwaχanás	<	*gʲə-wə-∅-χanás

These three basic forms are followed by a complexe set of morphemes, indicating tense, mood, and sometimes aktionsart. Often, the produced verb forms play the role of converbs. As the tense and/or mood inflected verbs do hardly differ structurally, it seems appropriate to discuss the paradigm altogether[93]. Normally, the gerunds and the infinitive are followed by a set of morphemes the semantics of which are heavily affected by the tense/aspect form of the verbal root itself. Still, the following generalizations can be made:

(81)	Verbal stem + ATTR.CM:	-na / -n	PARTICIPLE
	Verbal stem + SUFFIX:	-ni	CONFIRMED
		-j	PAST
		-xe	CONDITIONAL
		-wu	CAUSE
		-ji	UNCONFIRMED
		-oxu-CM	UNWITTNESSED
		-ɢˢaˢ	ASSUMED
		-let:i	EQUAL IN TIME
		-me(gə)	OPTATIVE
		-(j)kum	CONSECUTIVE
		-ʤar	SEPARATE ACTION[94]

[92] The fact that the CM series *-r-, -r-, -b-, -d-* appears in the past gerund (instead of *-∅-, -j-, -w-, -∅-*) is perhaps conditioned by the strongly lexical character of the morpheme *-wa-*. Also note its position *before* the CM which hints at a possible preverbial character of this element.

[93] The syncretism of TAM forms and participles and converbs is a typical phenomenon of the (S)ECL. It can be regarded as one of the basic isomorphic features in the area.

[94] In fact the CV *-ʤa-*CM has a strong contrastive connotation, from which its secondary function as a CV of negation is derived.

	-kan	CONDITIONAL CONFIRMED

Supplementary TAM forms are construed with the help of the auxiliary *wo*-CM (which itself appears only as a present gerund), and with case markers, which are added to the infinitive or to the masdar (verbal noun). The interaction of TAM suffixes, case markers, and auxiliaries with the stem forms of the verb produces the following TAM categories and converbs[95]:

(82)	**PRESENT GERUND**	**∅-e:x-e**	
	Participle	∅-e:x-e-na	(IV ∅-e:x-e-n)
	Confirmed	∅-e:x-e-ni	
	Confirmed imperfect (specific)	∅-e:x-e-ni-j	
	Confirmed imperfect (general)	∅-e:x-e-na-ni-j	(IV ∅-e:x-e-jn-ni-j)
	Unconfirmed	∅-e:x-e-ji	
	Conditional	∅-e:x-e-xe	
	Conditional imperfect	∅-e:x-e-xi-j	
	Iterative	∅-e:x-e-na-xe	(IV ∅-e:x-e-n-xe)
	Iterative optative	∅-e:x-e-na-xi-j	(IV ∅-e:x-e-n-xi-j)
	Unwittnessed	∅-e:x-e-na-oxu-CM	
	Assumed	∅-e:x-e-n-ɢˢaˢ	
	Equal in time	∅-e:x-e-let:i	
	Optative	∅-e:x-e-me:	
	Strong optative	∅-e:x-e-me:-gə	
	Separate action	∅-e:x-e-ʤar	
	Causal	∅-e:x-e-wu	
	Narrative	∅-e:x-e-wo-CM	
	PAST GERUND:	**∅-əx-a**	
	Participle	∅-əx-a-ni	(IV ∅-əx-a-n)
	Confirmed	∅-əx-a-jni	
	Confirmed past (specific)	∅-əx-a-nij	
	Confirmed past (general)	∅-əx-a-nan-nij	(IV ∅-əx-a-jn-ni-j)
	Unconfirmed	∅-əx-a-ji	
	Conditional	∅-əx-a-xe	
	Conditional past	∅-əx-a-xi-j	
	Iterative	∅-əx-a-na-xi-j-ni	(IV ∅-əx-a-jn-xi-j-ni)
	Unwittnessed	∅-əx-a-oxu-CM	
	Assumed	∅-əx-a-jn-ɢˢaˢ	

[95] As an example the verb *ixes* "to be(come)" has been taken. Only class I forms are mentioned, as the other classes can easily be derived from them with the help of the informations given above. Also note that the suffix *-nij* is restricted to first person agents (else *-ij* is used). When an attributive CM appears, the appropriate class IV marker will be indicated.

Equal in time	Ø-əx-a-letːi	
Optative	Ø-əx-a-meː	
Strong optative	Ø-əx-a-meː-gə	
Consecutive	Ø-əx-a-(j)kum	
Causal	Ø-əx-a-wu	
Narrative	Ø-əx-a-wo-CM	
Narrative past	Ø-əx-a-CM-a	
MASDAR:	**Ø-əx-aj**	
Comparative	Ø-əx-aj-le	
Dependent	Ø-əx-ajkʷan	
CV causal	Ø-əx-aj-ke	
CV causal confirmed	Ø-əx-aj-qa	
CV causal conditional	Ø-əx-aj-lqa	
CV causal (reason)	Ø-əx-aj-lʲ-alla	
CV inferential	Ø-əx-aj-na	(IV Ø-əx-aj-n)
Participle (confirmed)	Ø-əx-aj-ni	
CV telic	Ø-əx-aj-s	
FUTURE (INFINITIVE):	**Ø-ix-es**	
General future	Ø-ix-es	
CV causal	Ø-ix-es-wu	
Near future	Ø-ix-es-se	
Participle	Ø-ix-es-da	(IV Ø-ix-es-ən)
Confirmed	Ø-ix-es-di	
Future past I	Ø-ix-es-da-nij	(IV Ø-ix-es-ən-nij)
Future past II	Ø-ix-es-nij	
Conditional	Ø-ix-es-xe	
Conjunctive	Ø-ix-es-ij	
Conditional confirmed	Ø-ix-es-kan	
Conditional unconfirmed I	Ø-ix-es-da-(a)n-gˢaˢ	(IV Ø-ix-es-n-an-gˢaˢ)
Conditional unconfirmed II	Ø-ix-eskan-njij	
Future dubitative	Ø-ix-es-oxu-CM	
Optative	Ø-ix-es-meː	
Strong optative	Ø-ix-es-meː-gə	
Separate action	Ø-ix-es-ʤar	
Narrative	Ø-ix-es-wo-CM	

The semantic description of the above mentioned categories is only tentative. In many cases it depends on the idiosyncrasy of the speaker whether a specific TAM form is preferred. Thus the paradigm given in (82) can be regarded as the optimal system of TAM inflection in Tsakhur, incorporating both highly frequent forms (like the narratives) and rarely used, but still

productive categories (like the unconfirmed future conditional II (best to be translated by "if he (perhaps) will have done this or that...")).

The system of TAM formation in Tsakhur is further complicated by the fact, that some verbs show reduplication especially in the durative aspect, i.e. in present and future TAM forms. The technique of reduplication is somewhat unpredictable, it may affect the first consonant of the root as well as the whole root[96]. (83) gives an example for the verbs "to beat" and "to give birth"[97]:

(83)	"to beat":	MASDAR:	əˢχˢ-əˢj
		INFINITIVE:	əˢχˢ-aˢs
		PAST:	əˢχˢ-əˢ
		PRESENT:	əˢχˢ-əˢ'-χˢaˢ
	"to give birth":	MASDAR:	uχ-uj
		INFINITIVE:	uχ-as
		PAST:	uχ-u
		PRESENT:	uχ-oχ-a

4.8.5 Personal inflection

Most SECL do not mark verbs for speech act participants[98]. If agreement is present, this is done with the help of class markers. As Aghul and Lezgi lack class agreement, too, their verbal paradigms are highly isolating, cf. the Aghul (Richa) example:

(84) zun iqʷ'-a-j-a "I sit"
(I.ABS sit-TV.PRES-GERUND.PRES-AUX.PRES)

wun iqʷ'aja	"you sit"
mi iqʷ'aja	"he/she/it (here) sits"
xin iqʷ'aja	"we (incl.) sit"
ʧin iqʷ'aja	"we (excl.) sit"
ʧun iqʷ'aja	"you (pl.) sit"
mur iqʷ'aja	"they (here) sit"

The indication of persons on the verb is quite marginal in most of the SECL[99]. Sometimes it appears with imperative forms (e.g. in Khinalug and in Budukh). In Tsakhur we can observe a

[96] Reduplication in the SECL still remains an unstudied matter.
[97] Ibragimov 1990:133 limits reduplication to the present gerund, to the negative indicative, and to the prohibitive. There is, however, clear evidence, that reduplication also occurs with future TAM forms. cf. *-ək'ar-* "to be ill": *ʃena Ø-éːk'ar-(w)-or* (Present narrative) "he is ill (he says)", but *ʃena Ø-ək'-ejk'ar-ás-(w)o-r* (Future narrative) "he will be ill (he says)".
[98] The "unpersonal" character of the East Caucasian verb is often treated as one of the basic typological features of this language group.
[99] Udi is the only SECL that has developed a complete system of personal agreement also incorporating the non-personne, i.e. the third person. The SAP markers are clearly derived from the corresponding per-

tendency to separate the first person from the rest of the paradigm, reflecting a strategy to mark the central speech act participant out of pragmatic reasons[100]. The underlying technique is based on the use of the (determinating) attributive element *-na /-n* (cf. 4.3), which focuses the first person, cf. the paradigm of the auxiliary *wo*-CM ("to be", present tense):

(85)

	I	II	III	IV
1	zə wo-r-na	zə wo-r-na	zə wo-b-na	zə wo-b-on
2	ʁu wo-r-or	ʁu wo-r-or	ʁu wo-b-ob	ʁu wo-d-o-d
3	ʃena wo-r-or	ʃena wo-r-or	ʃena wo-b-ob	ʃen wo-d-od

Obviously, the attributive marker *-an /-n* replaces the final tense morpheme *-o*-CM (< *-*wo*-CM) in the first person (plural *-n*). Consequently, this technique can only be applied with other verbs, if the TAM form is based on this auxiliary, i.e. the narrative forms. As these forms are highly frequent, it can be assumed that person marking will start from them and will spread over the whole paradigm some days. (86) contrasts the person marked (past) narrative verb form Ø-əx-*a*-*wo*-CM with its unconfirmed past (Ø-*əx-a-ji*) of the verb *ixes* "to be(come)":

(86)

	I	II	III	IV[101]
1	zə Ø-əx-a-na	zə j-ix-a-na	zə Ø-ux-a-na	zə Ø-əx-a-n
2	ʁu Ø-əx-a-wo-r	ʁu j-ix-a-wo-r	ʁu w-ux-a-wo-b	ʁu Ø-əx-a-wo-d
3	ʃena Ø-əx-a-wo-r	ʃena j-ix-a-wo-r	ʃena w-ux-a-wo-b	ʃena Ø-əx-a-wo-d

	I	II	III	IV
1	zə Ø-əx-ji	zə j-ix-a-ji	zə w-ux-a-ji	zə Ø-əx-a-ji
2	ʁu Ø-əx-a-ji	ʁu j-ix-a-ji	ʁu w-ux-a-ji	ʁu Ø-əx-a-ji
3	ʃena Ø-əxa-ji	ʃena j-ix-a-ji	ʃena w-ux-a-ji	ʃena Ø-əx-a-ji

The syntax of personal agreement will be discussed in short in chapter 5.2.3. It is difficult to decide whether this type of personal agreement reflects only an actual pattern that pragmatically marks the SAP(1), or whether it is the starting point for a fully inflected system. The tendency towards an *egocentric* marking technique itself (opposing the SAP(1) to all other persons) in well known in ECL.

sonal pronoun, whereas the third person morphemes are shortened demonstratives. Tabasaran, too, has developed a system of (partly poly)personal agreement, though it is limited to speech act participants (i.e. to the first and second person). The *non-personne* is left unmarked iconically. Cf. Schulze 1997b for the whole question (with bibliography).

[100] Perhaps, this strategy has to be regarded as an areal feature: It is also shared by the Zakatal dialect of Awar and - what is perhaps more important - by the isolated dialect of Kusur (Awar), spoken in the northwest of Caxur.

[101] The forms with first and second person and class IV are naturally "odd", as referents of class IV can hardly function as speech act participants. Still, my informants accepted them saying that they could be used with anthropomorphic referents (for instance *zə (aslan) Ø-əx-an* "I (the lion) was....").

4.8.6 Participles and converbs

As it has been said above, participles and converbs can hardly be discriminated from TAM forms in Tsakhur. In fact, any TAM form can be used in subordinated clauses. Thus, no further informations are given here.

4.8.7 Infinitive and Masdar

The overall technique in the SECL to mark (telic) infinitives is to add the (PL) dative *-Vs* to the verbal stem[102]. This technique has also survived in Udi which has lost the morpheme in case inflection (cf. *tad-es* "to give", but *adamar-a* "to the man"). In Tsakhur like in other SECL the infinitive is used to form future TAM paradigms (see above). The morphemes to mark verbal nouns (masdar) can hardly be traced back to the proto-language, in Tsakhur - as we have seen - the element *-Vj* is used to indicate masdars. It is then inflected as an ordinary noun and forms the basis for converbs, see above.

5. Syntax[103]

5.1 Structure of the NP

The basic order within NPs in the SECL is determinans - determinatum, i.e. a dependent element precede its head. If class marking is possible, it appears with any sort of attribute (cf. 4.3). In Tsakhur, the following order of elements in NPs is observed:

(87) mana qˤ'oˤj-re j-ik'ar-na xunaʃːe
that(II).ABS two-II II-ill-II woman(II)
"Those two ill women"

qˤ'oˤ-ni w-uk'ar-ni balkan-a-n k'ar-ən ɢelʲ-bə
two-OBL III-ill-OBL horse(III)-SA-GEN.IV black-ATTR.ABS.PL leg(IV)-PL
"The black legs of the two ill horses"

As participles are treated like adjectives, they appear in the same position, forming subordinate relative clauses, cf.

(88) tʃoʤ-us aliwʃes-da balkan za-k'le heːge
brother-DAT buy.INF-PART.III horse(III).ABS I-AFF show.IMP
"Show me the horse that you will buy for your brother!"

[102] Khinalug and the Southern Samur languages Kryz and Budukh have developed new forms of the infinitive.

[103] It is not intended to give either a complete description of Tsakhur syntax nor its comparative treatment with respect to the other Lezgian languages. In fact, a comparative syntax of the SECL is still urgently wanted.

zə qˤaˤdaˤqːˤ-aˤn kitab maraχlə-da wo-d[104]
I.ABS read-PART.IV book(IV).ABS interesting-PART AUX.PRES-IV
"The book I'm reading is interesting."

Possessives too take the position of attributes and hardly differ from them structurally (cf. 4.3). This basic and unmarked ordering of NP constituents cannot be changed within NPs. The only elements that can follow a nominal head are postpositions (cf. 4.6).

5.2 Structure of the clause

5.2.1 Word order and attention flow

Word order in Tsakhur follows the expected accusative routines established by the appropriate type of attention flow. Thus a sentence normally starts with an agent (if present) or an addressee (with verba sentiendi). This place can (tough rarely) be occupied by other NPs (esp. locatives) in order to topicalize them. This syntactic strategy is contrasted by the fact that Tsakhur - as it had already been said - like most other SECL operates on the basis of an ergative syntax, cf.:

(89) balkan q'iwk'u
horse(III).ABS III.die.PAST
"The horse died."

dakːʲ-eː balkan aliwʃu
father(I)-ERG$_{hum}$ horse(III).ABS III.buy.PAST
"The father bought a horse."

The absolutive case is used to mark intransitive agents and transitive patients whereas the transitive agent is encoded with the help of the ergative case. It is, however, difficult to tell whether these case forms indicate semantic roles or the syntactic functions subject and object. Tsakhur like most other SECL meet the conditions of morphological ergativity by using the appropriate case markers. The system of agreement discussed below hints at the fact that the view point starts from the transitive patient, that is, the patient plays the role of the subject, the ergative marked agent that of the object. This "patient-oriented" type of ergativity, however, involves (at least in the Tsakhur dialect of Tsakhur) - as it had been said - some aspects of accusativity. The most important split is invoked by the fact that the phrasal attention flow starts with the transitive agent and not with the patient (as expected for syntactic ergativity[105]). Thus the basic word order is more or less European-like, cf.:

[104] But cf. *zə kitab qˤaˤdaˤqˤaˤn maraχləda wod* "It is interesting that I read I book."
[105] The most common example for the matching of the (ergative) view point (centered on the transitive patient) with the attention flow is given by Dyirbal, cf. Dixon 1994 for details.

(90) dakːʲ-eː χaw alʲaʔu
father(I)-ERG$_{hum}$ house(IV).ABS IV.build.PAST
"The father built a house."

*(?) χaw dakːʲ-eː alʲaʔu[106]

The second position in a simple sentence is taken either by the verb or (if present) by a second NP. There seems to be no perceivable difference between a NP-V-NP and NP-NP-V ordering at least on the syntactic level, cf.

(91) jizin do t'ak'u wodun ~ jizin do wodun t'ak'u
my.IV name(IV).ABS Taku IV.AUX my.IV name(IV) IV.AUX Taku
"My name is Taku." "Dto."

However, a verb final structure is generally preferred in pragmatically unmarked sentences. Locatives, addressees, and adverbs are normally placed before a transitive patient (if present). Thus the overall scheme would be:

(92) AGENT - LOCATIVE - ADDRESSEE - PATIENT - VERB

An example (not always excepted by my informants) would be:

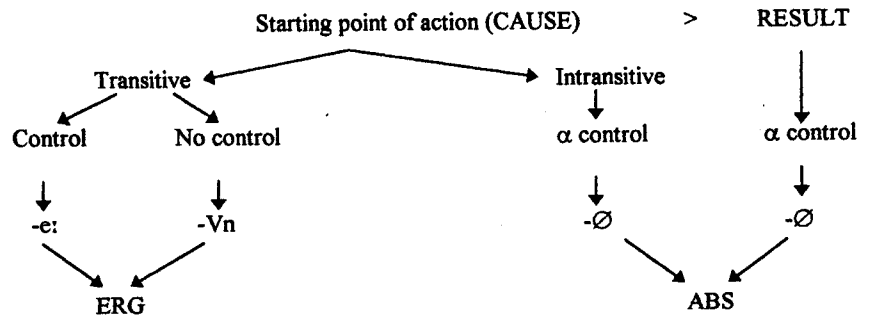
(93) dakːʲ-eː sanəxa χaj-a hajna kaʁəz ojk'an-ij
father-ERG$_{hum}$ yesterday house-IN.ESS one-III letter(III).ABS III.write.PAST-FOC
"Father has written this letter at home yesterday."

5.2.2 Semantic roles

In general the semantic (hyper)roles AGENT and PATIENT are encoded on the basis of an ergative coding strategy. However, it must be noticed that - as it has been said - the ergative case itself is split according to the dichotomy [±hum] (*-eː* vs. *-Vn*) in the singular. In fact, only referents [+hum] are thought to have a true agentive potential, whereas non-humans are seen as entities without an inherent feature of agentivity. The aspect of agentivity or of control seems to be relevant only when referring to transitive states of affairs. Accordingly, the morphemes *-eː* and *-Vn* can be subsumed under a label "starting point of action towards another entity" (CAUSE). The intransitive "agent" itself is left unmarked, because it lacks this orientation ([αcontrol]). It is treated just like the entity that lies in the RESULT (or EFFECT) domain of transitive actions, cf.

[106] My informants have generally denied the possibility of placing the patient at the beginning of the proposition in an unmarked context. A rough glance through the data of the other SECL gives a similar impression, though the rigidness of placing the (in)transitive agent in front of the verb may vary from language to language.

(94)



Cf. the following two sentences, the first of which denotes that the man has destroyed the bridge on purpose, whereas in the second sentence the river "does" it hazardously:

(95) adam-eː jiˢʁˢ alebt'ə
man-ERG$_{hum}$ bridge(III).ABS III.destroy.PAST
"The man destroyed the bridge."

dama-n jiˢʁˢ alebt'ə
river-ERG$_{-hum}$ bridge(III).ABS III.destroy.PAST
"The river destroyed the bridge."

Contrary to some other SECL Tsakhur hardly knows any means to encode the transitive agent other than by the ergative case. An example for such variations is given by Archi, cf. the two following sentences, which differ in the degree of affectedness and reference[107]:

(96) q'uˢt'i-li lo eˢwq'ni
thunder-ERG boy(III).ABS III.frighten.AOR
"The thunder frightened the boy."

q'uˢt'i-li-t͡ɬ:'iʃ lo eˢwq'ni
thunder-SA-SUB.ABL boy(III).ABS III.frighten.AOR
"The boy was afraid of the thunder."

One exception is the verb *aˢχaˢs* "to be able", which has its agent in the AD.ABL, cf.

(97) dekːʲ-isːe aˢχaˢn qaˢtqaˢs kitab
father-AD.ABL can.PRES IV.read.INF book(IV).ABS
"Father can read the book."

[107] Alekseev 1979:87. This is the example for only one type of variation in Archi. The language offers a great range of possibilities to reduce the (assumed) agentivity of an agent by changing case marking (and hence valence).

The transitive patient is always unmarked with respect to case forms. There are no means to subcategorize the patient (e.g. like Udi ([±definite]). Hence, it comes clear that (with non-SAP referents) Tsakhur exhibits a relatively strong patient oriented system which allows to assign the syntactic function "subject" at least partly to the transitive patient, whereas the ergative case serves to denote the object. This ergative typology is again split by another aspect of accusativity which is introduced by the fact that the Tsakhur dialect (not Gelmets) today follows some of the predictions of the well-known Silverstein Hierarchy: With agentive speech act participants the case marking is at least "non ergative" (though agreement still is), cf.:

(98) ʁu χaː-qa qa-r-ə
you(I).ABS house-ALL return-I-PAST
"You returned home."

ʁu balkan gʲoːtu-wo-b
I(I).ABS horse(III).ABS III.beat-PAST.NARR-III
"You have beaten the horse."

xunaʃː-eː ʁu gʲetu-wo-r
woman(II)-ERG$_{hum}$ you(I).ABS I.beat-PAST.NARR-I
"The woman has beaten you."

The accusativity is still strengthened if a first person is involved as an agent, as now the verb (in single tense forms, cf. 4.8.4) also agrees with the pronoun, cf.[108]:

(99)

1.SG$_{AG}$:	2.SG$_{AG}$:
zə deχ gʲetu-na I(I).ABS son(I).ABS I.beat.PAST-SAP(1) I have beaten the son."	[ʁu deχ] gʲetu-wo-r I.beat.PAST-NARR-I"
zə jiʃ gʲeːtu-na I(I).ABS girl(II)ABS II.beat.PAST-SAP(1) "I have beaten the girl."	[ʁu jiʃ] gʲeːtu-wo-r II.beat.PAST-NARR-II
zə balkan gʲoːtu-na I(I).ABS horse(III).ABS III.beat.PAST-SAP(1) "I have beaten the horse."	[ʁu balkan] gʲoːtu-wo-b III.beat.PAST-NARR-III
zə aslan gʲetu-jn I(I).ABS lion(IV).ABS IV.beat.PAST-SAP(1$_{IV}$) "I have beaten the lion."	[ʁu aslan] gʲetu-wo-d IV.beat.PAST-NARR-IV

However, the Silverstein Hierarchy does not operate in all SECL, and probably did not affect case marking in PL. Thus case syncretism with personal pronouns should be regarded as a

[108] The glossing of the complete phrases with second person agent is neglected as it can easily be construed.

natural strategy that may be observed by single languages but not by the whole language group or by PL. This also comes clear from Gelmets which not only has preserved the old ergative morphology of personal pronouns (*zə* vs. *za-*, *ʁu* vs. *wa-* etc., cf. (54)), but which has even strengthened this morphology by introducing the innovated ergative morpheme *-s:a*.

5.2.3 Agreement

The agreement pattern - as it has been said above - is basically ergative, though aspects of accusativity appear with pronominal agents. The function of verbal class markers is to indicate the syntactic center of a phrase, i.e. the absolutive marked NP. The centrality of this NP is additionally marked by the fact that it can serve as a agreement trigger for nearly every other constituent in a sentence (see below). A secondary aspect of agreement is introduced in Tsakhur by the expansion of (nominal) plural markers to verbs, especially with intransitive agents or transitive patients, cf.:

(100) ʃi tʃ'alaga-qa w-uˢqːˢaˢ-n-bə
we(I/II).ABS wood-ALL I/II.PL-go.PRES-SAP(1)$_{PL}$-PL
"We go into the woods."

zə itʃij gjeːt'u-na
I(I).ABS girl(II).ABS II.hit.PAST-SAP(1)
"I hit the girl."

zə itʃ-er gʲeːt'i-jn-bə [109]
I(I).ABS girl-PL.ABS II.hit.PAST-SAP(1)$_{PL}$-PL
"I hit the girls."

zaz jizən jitʃij-bə b-əkːanan-bə
I.DAT my.II.PL sister(II).ABS-PL II.PL-love.PRES-PL
"I love my sisters."

This type of plural marking can occur even if a open NP is not present, cf. (Gelmets):

(101) ʤes-qa abajen-bi məslaˢʁˢ-aˢt-əz
REFL-ALL I/II.PL.come.PRES-PL advice-AUX-INF
"(They) come to HER for advice."

[109] But not **ʃi itʃij gʲeːtena-bə* "we hit the girl" etc., a hypothetical form of agent agreement in Tsakhur that sometimes is allowed in the neighboring Awar dialect of Kusur.

5.3 Major sentence types

As had been said above the ergative coding strategy dominates nearly every sentence type in Tsakhur. Consequently the dichotomy [±transitive] plays an important role. Tsakhur is not a good example for other processes of syntactic derivation in the SECL, as passives, antipassives, pseudo-transitives, labile constructions, etc., which can occur some of the other languages. In fact, Tsakhur lacks all these procedures, it is much more "role dominated" than e.g. Udi[110]. Causatives are marked analytically with the help of the transitive auxiliary *haʔas* "to make", added to the infinitive. The embedded agent is either demoted to an absolutive or maintained as an ergative, cf.

(102) jiʃːe kaʁəz ojk'an-ij
daughter-ERG$_{hum}$ letter(III).ABS III.write.PAST-FOC
"The daughter wrote a letter."

dakːʲ-eː jiʃːe: (~ jiʃ) kaʁəz ojk'as haːʔa ~ hajʔa
father-ERG$_{hum}$ daughter-ERG$_{hum}$ (~ ABS) letter(III).ABS
write.INF III.make.PAST ~ II.make.PAST
"Father had (his) daughter write a letter."

One major exception to the stableness of ergative marking is the well-known dative or "inverse" construction which is a typical feature of all East Caucasian languages. With verbs of perception and affection the experiencer is encoded by the dative and/or a locative case form, whereas the (absolutive) patient still is the agreement trigger. In Tsakhur we have two different types of verba sentiendi: Verbs of affection have their experiencer in the dative, whereas verbs of perception show the affective case form[111], cf.:

(103) dakː-is deχ ∅-əkːan
father(I)-DAT son(I).ABS I-love.PAST
"The father loved the son."

dakː-is-əd pəl laːzim-da wo-d-on
father(I)-DAT-IV money(IV).ABS need-ATTR AUX-IV-PRES.IV
"The FATHER needs money."

[110] In this respect, Tsakhur behaves much like Khinalug, which is relatively role dominated as well(cf. Schulze(-Fürhoff) 1994a). True antipassives are not attested in SECL, passives only in Udi and (perhaps) in Kryz. Analytic procedures such as "binominatives", which foreground the agent of a transitive phrase but do not background the patient are known from Archi, Rutul, and Khinalug (cf. fn.17). Practically, all processes related to the strategy of grounding have to be described in terms of single language grammars, if they are present at all.

[111] The affective case is used to encode the addressee with the verbs *ehes* "to say" and *hagʷas* "to show", cf. *zə direktor-ik'le-d kitab hagu-jn* (I.ABS director-AFF-IV book(IV).ABS IV.show-PAST.NARR.SAP(1)$_{IV}$) "I have shown the book to the director$_{TOP}$").

Mansur-e-k'le jiʃ ɢʲiʤ-esda
Mansur-SA-AFF girl(II).ABS II.see-FUT
"Mansur will see the girl."

jiʃ-e-k'le Mansur ɢaʤ-esda
girl(II)-SA-AFF Mansur.ABS I.see-FUT
"The girl will see Mansur."

za-k'le ∅-aʦa deʃ
I-AFF IV-know.PRES NEG
"I do not know (it)."

The dative case also serves as the normal means to encode any sort of addressee in all SECL. Contrary e.g. to Archi the dative is rarely involved in class agreement in Tsakhur (except for focus, see below), cf. as an example:

(104) dakːʲ-eː duχ-aj-s maʃin aliwʃu
father-ERG$_{hum}$ son-SA-DAT car(III).ABS III.buy.PAST
"The father bought [his] son a car."

Q'urban-ni-d ʧoʤ-us kitab hiwu
Qurban-GEN$_{OBL}$-IV brother(I)-DAT book(IV).ABS IV.give.PAST
"He gave the book to QURBAN's brother."

In general, reflexivization as a syntactic procedure is quite marginal in most SECL. The use of reflexive pronouns in patient position is found only in some languages, cf. e.g. the Budukh examples[112]:

(105) gədə-z ug-ug ɨrqaʤ-i
boy(I)-DAT REFL-REFL.ABS see-PRÄT
"The boy saw himself."

riʒ-əz gədə-rə ug naˤrə-syʔy laːzim i
girl(II)-DAT boy(I)-ERG REFL.ABS call-make.INF necessary AUX.III
"The girl must be called from her brother [lit. "her boy]."

In Tsakhur, the reflexive pronoun is basically used to emphasize a subject NP and appear very often as substitutes for the possessive (genitive) pronouns. Sometimes, the reflexive pronoun can even replace the non emphatic demonstratives used to indicate a ¬SAP, cf.

(106) ʤo adam-er
REFL(I).PL.ABS man(I)-PL.ABS
"The men themselves..."

[112] Alekseev 1994b:268.

jiʤ-bə balkan-ar
REFL(III)-PL.ABS horse(III)-PL.ABS
"The horses themselves..."

(107) jiʃna ʁuʒ ki jugna-j ki jamanə-j wor-na[113]
our.I REFL EMPH good-and EMPH bold-and I.AUX.PRES
"He$_{TOP}$ (which is our [father]) is very good and very bold."

If the reflexive is use attributively (in the sense of "his/her/its own...") the normal trigger operates on an accusative scheme, i.e. it nearly always is the agent of a sentence, cf. the example in (114). However, such a type of reflexivization is quite marginal.

5.4 Pragmatic functions

In many cases agreement is extended to focus functions: Any noun, adverb, or (partly) postpositional phrase can be class marked in order to be focused or topicalized[114]. Obviously the focus function then has been transferred from the absolutive marked noun (which is in topic function in unmarked positions) to the noun, adverb etc. agreeing with it. Though there are practically no limits as for which constituent can be in focus, the first NP or adverb of a phrase, i.e. the starting point of the attention flow is the preferred target of focus by class agreement, cf.:

(108) deχ-ər qˤaˤdqˤaˤs ark'ən
son(I).ABS-I learn.INF I.go.PAST
"The SON went to learn."

iʧij-r χaː-qa qa-r-ə
girl(II).ABS-II house-ALL return-II-PAST
"The GIRL returned home."

dakːʲ-is-ər jed-is-ər ʧu-bi-ʃi-s-ər, jiʧe-bi-ʃi-s-ər zə ∅-əkːan-o-r-na
father(I)-DAT-I mother-DAT-I brother-PL-SA-DAT-I sister-PL-SA-DAT-I
I(I).ABS I-love-NARR-I-SAP(1)
"[My] FATHER, MOTHER, BROTHERS, and SISTERS loved me."

(109) dakːʲ-eː-r jiʃ qˤaˤdqˤaˤs ɢaxo
father(I)-ERG$_{hum}$-II daughter(II).ABS learn.INF II.send.PAST
"FATHER sent [his] daughter to learn."

[113] This is a Gelmets example.
[114] This technique is unique within the Samur languages, but partly attested for Archi (e.g. adverbs and addressees can be class marked here).

Another technique of focusing or topicalizing constituents is the use of the particle *-(i)j* (*-nij* with first person participants). As for position this element is much more free than agreement markers, cf.:

(110) zə kaʁəz ojk'an-n-ij
I(I).ABS letter(III).ABS III.write.PAST-SAP(1)-TOP
"I wrote$_{FOC}$ a letter."

ʁu kaʁəz ojk'an-ij
you(I).ABS letter(III).ABS III.write.PAST-TOP
"You wrote$_{FOC}$ a letter."

ʃen-ɢ^{w}-eː kaʁəz ojk'an-ij
he-SA$_{m}$-ERG letter(III).ABS III.write.PAST-TOP
"He wrote$_{FOC}$ a letter."

The following focus variants can be observed:

(111) 1.SG zə kaʁəz$_{FOC}$-nij ojk'an, zə$_{FOC}$-nij kaʁəz ojk'an
2.SG ʁu kaʁəz$_{FOC}$-ij ojk'an, ʁu$_{FOC}$-ij kaʁəz ojk'an
3.SG ʃenɢ^{w}e kaʁəz$_{FOC}$-ij ojk'an, ʃenɢ^{w}e$_{FOC}$-jij kaʁəz ojk'an

Other particles used for emphasis or focusing are e.g. *bes* (placed after any word in a phrase[115]), *le* (in the sense of "but"), *mer* (I,II) / *meb* (III) / *med* (IV) (indicates an iterative action), *-maː* ~ *-meː* (emphasis), cf.

(112) ʃen-bə me-b tʃalaga-qa habk'ən-ob
DEM-PL again-I/II.PL wood-ALL go.PAST-AUX.I/II.PL
"They again went into the wood."

zə bes kaʁəz ojk'an-nij
I.ABS EMPH letter(III).ABS III.write.PAST-FOC.1.SG
"I wrote the letter."

zə t̪s'aˢχa-qa-maː qa-r-ə
I.ABS Caxur-ALL-EMPH come-I-PAST
"I came to Caxur."

In fact, *bes* can be used as a simple affirmative particle is the sense of "yes", cf.

(113) bes bes hˢaˢmaˢxˢuˢd heʔe
EMPH EMPH so make.IMP
"Yes, yes, do it this way!"

[115] Like many other particles and junctions, Tsakhur *bes* is a loan (< pers. *bas*, originally a particle used to indicate limitation).

5.5 Negation

Negative clauses follow the same word order as affirmatives, because negation is basically bound on verb morphology. Tsakhur uses the prefix *-de-* with variants, which is placed after a CM (if present). Thus we have *∅-əxa* "was (I)" vs. *de-∅-xa* "was not (I)", *j-ixa* "was (II)" vs. *jə-də-xa* "was not (II). In normal speech an analytical construction is preferred using the negative auxiliary *deʃ* (restricted, however, to the present, perfect, and future tense forms). Prohibitives are construed with the help of the prefix *m-* ~ *mV-* added to present gerunds exclusively. Its position with the prefix chain is that of the ordinary negation *de-* (cf. above). The scopus of negation can be clarified with the help of negative pronouns, e.g. *haʃːu-ʤa*-CM "nobody", *hiʤo-ʤa*-CM "nothing"[116], *hitʃːub* "nothing", cf.:

(114) ʃawaʤar ʤuna dost ɢaleratʃːe-na deʃ
nobody.ERG REFL.GEN friend(I).ABS I.abandon-PART NEG.AUX
"Nobody abandons his friend."

ʃawniʤar χaː imaχ^{w}a
nobody.GEN house.LOC stay.PROH
"Do not stay in anybody's house!"

niʃeʤar ajt̪s'e-na deʃ[117]
nothing.ERG satiate-PART NEG.AUX
"A nothing does not satiate [anybody]."

5.6 Questions

Questions are formed either with the help of the Q-words *haʃːu* "who", *hiʤo* "what" (and their derivatives), or - for yes/no questions - by using the particles *-ne* / *-je* (cf. (58), for the distribution of *-ne* /*-je* see examples in (63)):

(115) kaʁəz ok'ne-ne
letter(III).ABS III.write.IMP-Q
"It there a letter to write$_{FOC}$?"

kaʁəz-ne ok'ne
letter(III)-Q III.write.IMP
"It there a letter$_{FOC}$ to write?"

[116] The element *-ʤa-* is clearly related to the suffix for "separate action" *-ʤar-*. In fact, it basic semantics is covered by something like "except" (cf. also *zə-ʤar* "except me" etc.).
[117] A proverb giving a rare example of patient deletion in a transitive sentence.

dʒamalij ∅-atʂ'a-χe-ne[118] t'ak'u bazara-qa uˢqˢaˢs hˢaˢzər-qa-je
Dshamaliy I-know-COND-Q Taku.ABS bazar-ALL go.INF ready-AUX-Q
"Dshamaliy, do you know whether Taku is ready to go to the bazar?"

5.7 Coordination of NPs and Clauses

Coordination of NPs and clauses in the SECL often is asyndetic, that is constituents are chained without any morphological means. Tsakhur uses particles like *ji* ~ *ij* "and", *wa* "and", *de* "or", if necessary.

The NP junction *-ij* ~ *-ji* either appears after the first NP or reduplicated after both of them, cf.

(116) dakːʲ-ij jedʲ ~ dakːʲ-ij jedʲ-ij
father-and mother — father-and mother-and
"Father and mother" — "Dto."

ʈʃoʒ-er-əj jiʈʃi-bi ~ ʈʃoʒ-er-əj jiʈʃi-bi-ij[119]
brother-PL-and sister-PL — brother-PL-and sister-PL-and
"brothers and sisters" — "Dto."

The sentence junction *wa* is not as common as in some other SECL. Especially in folk tales the asyndetic structure is preferred. However, in conversation *wa* often indicates that the speaker intends to add another sentence to his talk, but that he is still unsure about it. The particle is also used to link to adverbs or predicative adjectives, cf.

(117) man-bi-ʃ-eː iʃ jug-ba wa ək'-ba haːʔa
DEM-PL-SA-ERG work(III).ABS well-III and quickly-III III.do.PAST
"They did (their) work well und quickly."

5.8 Subordination

Subordination is very limited in the SECL (except for Udi). Instead, converbs, infinitives, and participles are used. Examples have already been given throughout the text, so that only some additional data are given here. Participles are used to encode relative clauses, cf.:

[118] Ibragimov 1990:136 has *atʂ'axeeːle* instead of *atʂ'aχene* in a quite parallel example. However, such a form never occurred in my notes. What we have is a form *atʂ'aχe-le* which is a conditional followed by an element *-le* denoting something like "so that", cf. *ilekːe-le hima uftanna iʈʃije* "look, what a nice girl!" (lit. "look (so that I can ask you:) is she a nice girl?" (cf. "Look! Isn't she a nice girl?")).

[119] Again a Gelmets form. Tsakhur would have *ʈʃu-ba-ij jiʈʃi-bi(-ij)*.

(118) zə ʈʃodʒ-u-k'le gagu-jn kitab maraχlə-da wo-d
I(I).ABS brother(I)-AFF show-PART$_{IV}$ interesting-ATTR AUX-IV
"The book that I have shown [my] brother is interesting."

jiʃ wudʒ-eː dakːʲ-eː alʲaʔu-jn χaː-qa qa-r-ə
girl(II).ABS REFL(I)-ERG$_{hum}$ father-ERG$_{hum}$ build-PART.PAST$_{IV}$
house(IV)-ALL return-II-PAST
"The girl returned to the house that her father had built."

hajna zə aliwʃu-na balkan deʃ
this(III).ABS I(I).ABS buy-PART.PAST-ATTR$_{III}$ horse(III).ABS NEG.AUX
"This is not the horse that I have bought."

zə ∅-eːxe-ni χaː paprəs ʈsəˢʈsəˢaʔan deʃ
I(I).ABS I-be-PRES.CONF room.LOC cigarette smoke-PART.PL NEG.AUX
"In the room in which I am they don't smoke."

t'et' deʃ-di baʁ-eː bulʲbulʲ hoːk'a-na deʃ
flower.ABS NEG.AUX-ATTR$_{OBL}$ garden-LOC nightingale(III).ABS
III.sing-PRES NEG.AUX
"In a garden without flowers a nightingale does not sing."

Agreement patterns and case marking is also preserved with converbs and infinitives, often the nominal head (agent) of the superordinate clause is placed in the first position of the phrase, if it is coreferential with the agent of the embedded clause (i.e., cross referencing operates on an accusative pattern):

(119) χaj-bə alʲaʔas ɢaje ʃi dam-eːntʃe alqaʔu
house(IV)-PL.ABS build.INF stone.PL.ABS we river-IN.ABL IV.lift.PAST
"In order to build houses we took stones out of the river."

ʁu ixes-wu zə qa-r-ə
you(I).ABS I.be-there-CV(CAUS) I(I).ABS come-I-PAST
"I came because [I hoped that] you were there."

wa-sːe ixes-xeː heʔe
you(I)-AD.ESS be.INF-COND do.IMP
"If you can, do [it]."

balkan w-uk'ər-wu qiwk'u
horse(III).ABS III-ill-CV(CAUS) III.die.PAST
"Because the horse was ill it died."

Due to Azeri (and Russian) influence more and more subordinating particles enter the Tsakhur grammar. Moreover, Tsakhur tends to imitate their functionality by using specific suffixes (often in second position). Thus we have *gora* "because", *-xe* "if", *-kʷa* "how", *alla* "as" etc.

These elements often simply support the converbial form or serve as modal adverbs, chaining two sentences by coordination rather than subordination. An example would be:

(120) zə qˤaˤla-kʷa worna mantʃi-qa gora jiʃon-aʔa deʃ
I.ABS angry-how I.AUX.PRES DEM.SA-ALL because talk-AUX.PRES NEG
"I am angry, that is why I do not talk."

6. Sample text

The following text is taken from Ibragimov 1968:106[1-15]. The dialect is that of Tsakhur-Suwagil' (village of Mišleš), though "Tsakhurisms" can sometimes be observed (e.g. the ERG_{hum} *-e:* instead oft *-i:*) The text is slightly corrected with the help of informants. Due to the morphoponological processes typical for Tsakhur the glossing of many words cannot be done via a simple segmentation. Thus the heavy *portmanteau* structure especially of verb forms is maintained in the glosses. Most of the forms, however, can be analyzed with the help of the informations given in the appropriate chapters. In some cases an additional interpretation is given in the notes. The inherent class of a noun is indicated only if it is the trigger for any sort of agreement. Contrary to the examples given in the grammar accent is always indicated in this text.

The text itself demonstrates the typical way Tsakhur people talk about (historical) events in their villages. It is basically asyndetic. The informant switches from past to present tenses more than once in order to tell about the fate of a man from Mišleš, called Kawkha (*kawχá*)[120].

Kawχá miʃléʃ-ni χiw-é jedíke əx-á.
Kawkha.ABS Mishlesh-GEN(IV) village(IV)-IN/ESS born I.AUX-PAST
"Kawkha was born in the village of Mishlesh."

k'əlʲwál-eː dakːʲ qik'-ú jetírma aχ-ú.
youth-IN.ESS father.ABS die-PAST orphan(I).ABS I.become-PAST
"[His] father died, when he was young, [and] he became an orphan."

molʲ-ní jeʃ-é peʃakár əx-á.
eight-ATTR(IV) age(IV)-IN.ESS shepherd(I).ABS I.AUX-PAST
"By the age of eight he became a shepherd."

Kawχá jit̪s'ə́molʲe-ni [121] sen-eː́ bəkə́r-na adámij eːxʲ-e.
Kawkha.ABS nineteen-ATTR(IV) year(IV)-INESS adult-ATTR(I) man(I).ABS I.AUX-PRES
Kawkha was an adult, when he was nineteen years old."

[120] A much more complexe textual structure is given e.g. in the tale recorded by Dirr 1913:124-127. However, this text is full of errors. It has not been quoted here because I did not check it with informants.
[121] **jit̪s'-é-molʲi*-ATTR ("ten-plus-nine-"), cf. 4.5.

Kawχá-is Asːíjat-uka dawát-bə haʔ-á.
Kawkha-DAT Assiyat-SUB.ADESS marriage(III)-PL.ABS III.make-PAST
"Kawkha married Assiyat."

Kawχá waˤq'áˤ-bi-ʃi-kʷa q'əˤdím-əs ajk'án Galá-qa.
Kawkha.ABS sheep-PL-SA-COM winter-DAT I.go-PRES Gal-IN/ALL
"In winter time Kawkha goes with the flock of sheep to [the village of] Hala[122]."

Asːíjat χiw-é abáj-kʷa éːχʷ-a.
Assiyat village-IN.ESS mother=in=law(II)-COM II.stay-PAST
"Assiyat stayed in the village with [her] mother-in-law."

bak'istáw-ən duχ-jé gʲéːlles χiw-ə́-ni áːχʷ-a-m-mə [123].
Shepherd-GEN.PL son-PL.ABS must.INF village-IN.ESS-IV stay-PAST-OPT-PL
"The shepherd's sons had to stay in the village."
mám-mə me-b χiw-é-ni.[124]
DEM.ABS-PL again-I/II.PL village-IN.ESS-IV
"They were in the village again."

duχ-íʃ-da sa Asːíjat-ni qihná iːkár gʲiʁál.
son-PL-GEN(I) one.ABS Assiyat-GEN.OBL after I.go.GER PV.IV.begin.PRES[125]
"One of the sons pursues Assiyat."

sa jiʁə́lʲ Asːíjat t̪ʃawrá-jsqa j-iˤqːáˤn-ɢaˤ ɢa-t̪s'aːkʷ'-an jaˤq.
one day-SUPER Assiyat cattle-AD.ALL II-go-PRES-ASSUME
PV-III.block-PRES way.ABS
"One day, as Assiyat goes to the cattle, he blocks the way."

Asːíjat-ni xəˤwáˤ ajrqú gʲiːʁál ubá-bə haʔ-á.
Assiyat.ABS-ATTR.OBL hip(III).ABS III.take.PAST PV.IV.begin.PRES
kiss(III)-PL.ABS III.PL.make-PAST[126]
"He embraces Assiyat und gives her a kiss."

Asːíjat man-ɢú-s ʃillé ə́ˤwχ-əˤ héːχʷa-na.
Assiyat.ABS he-SA_m-DAT box(III).ABS III.give-PAST II.go=away-PRES
"Assiyat gives him a box on the ear [and] goes away."

[122] A village in Azerbajdzhan.
[123] Assimilated form of **áːχʷa-m-bə*, which itself has to be interpreted as**á-w-χʷ-a-meː-bə* (PV-I.PL-stay-PAST-OPT-PL).
[124] Literally: "They were in-village-beings".
[125] The verb *gʲiʁales* "to begin" often indicates an inchoative. Its class IV agreement refers to the whole embedded sentence.
[126] *haʔá* "made" contains the CM -∅ for PL III (**ha-∅-ʔ-a*). In the singular it has assimilated the infix **-b-* (*haːʔa* < **ha-w-ʔ-a*).

χa: abá-is χabár há:ʔ-a.
at=home mother=in=law-DAT news(III).ABS III.make-PAST
"At home she informs [her] mother-in-law."

ab-é [127] ejhén.
mother=in=law-ERG$_{hum}$ III.say.PRES
"The mother-in-law says:"

As:íjat! Kawχá-is negáhba χabár hi-ma-ʔ-a.
Assiyat Kawkha-DAT nothing news(III).ABS PV-PROH-make-PAST
"Assiyat, do not inform Kawkha!"

duχ-áj-k'le [128] ɢaj-xˤé eb w-úxʲ-es jíɣ-ən-ɢu-ni oˤgʲilʲ sán-ɢʷ-e: hidʒó haʔ-é:.
son-SA-AFF IV.know-COND blood(III).ABS III-be-FUT seven-SA$_m$-ATTR.PL
against one-SA$_m$-ERG$_{hum}$ what.ABS do-PRES.Q
"If [my] son knows [it], [there] will be blood, what can a single [person] do against seven!"

Kawχá aˤzázilʲ insán wó-r-na.
Kawkha.ABS rigorous man(I.ABS) AUX-I-PRES
"Kawkha is a rigorous man."

dʒan balá ses-de-ná gʲíjər-e [129].
love child voice-NEG-ATTR PV-II-AUX-IMP
"Dear child! Be silent!"

bak'é-qa abíj-ni ɢorʔána:-ʃ-e: Kawχá-is χabár há:ʔ-a.
sheep-ALL going-ATTR.OBL shepherd-PL-ERG Kawkha-DAT
news(III).ABS III.make-PAST.
"The shepherds going to the flock of sheep told Kawkha the news."

Kawχá dajánmiʃ de-x-á qájl-e χiw-é-qa.
Kawkha calm NEG-become-PRES I.return-PRES village-IN.ALL.
"Kawkha does not rest calm [and] returns to the village."

[127] As every noun showing the word formation suffix *-ij* (cf. 3.) *abaij* "mother-in-law" looses this element in the obliquus. The final *-a* is then contracted with the ergative morpheme *-e:*.
[128] A tsakhurism. Mishlesh would have *duχajk'li.*
[129] *-de-* is here used as an α-privativum. The resulting form then is used as an nominalized attribute ("the voiceless one"). The whole phrase would read: "Dear child, be you (CL II) a voiceless one!".

7. Bibliography

7.1 A brief look at the history of Tsakhur linguistics

As it has been said in the introduction linguistic work on Tsakhur started with the grammar of A. Dirr (Dirr 1913). Though it is full of mistakes and errors, it still contains some valuable information which should be once again checked with native speakers. The best reference book on Tsakhur surely is Ibragimov 1990. It describes *in extenso* the dialect of Tsakhur and also deals with the Gelmets dialect and two other (sub)dialects (Sabunchi and Suwagil). Due to the shortness of paper, however, this book is written in a very condensed form and not easy to use. The alternative would be ʒeiranišvili 1983/84, published in Georgian (with a Russian summary). Yet, this grammar does not reach the standard of Ibragimov's work. Moreover, it concentrates on the author's thesis that Rutul and Tsakhur are two dialects of a single language, which hinders him to give a mere descriptive treatment of these "dialects". There is another overview on Tsakhur by B.B. Talibov in the series *Jazyki narodov SSSR*, vol. 4 (Talibov 1967). Its fragmentary character does not allow its use as a primary source.

Some papers on Tsakhur can be found in the journals *Iberijsko-Kavkazskoe jazykoznanie* and *Ežegodnik iberijsko-kavkazskogo jazykoznanija* (mainly by Ibragimov and ʒeiranišvili), as well as in some collections of the Daghestan Institute of Languages and Literature (Makhatchkala), which today is the center for Tsakhur studies besides the Georgian Institute of Linguistics (Georgian Akademy of Science).

A dictionary is still urgently wanted. The only collection of words we have is given by Dirr 1913 (some 500 words). Kibrik/Kodzasov 1990 as well as Starostin/Nikolayev 1994 list a number of Tsakhur word for comparative reasons, however, the lexical information is quite marginal. We are also in need for a text collection. Dirr 1913 gives one longer text, one anecdote, and some verses, Ibragimov 1968 has one text and eight (shorter) poems, and Ibragimov 1990 offers one poem, some proverbs, and one shorter text (from Gel'mec). Perhaps new material will be accessible when publishing in Tsakhur starts on basis of the new written language.

7.2 References

The bibliography mentions only those titles which were referred to in the text.

ALEKSEEV, M.E. 1979. Funkcii èrgativnogo padeža v arčinskom jazyke. *Imennoe sklonenie v dagestanskich jazykax* (otv. red. U.A. Mejlanova). Maxačkala: IIJaL, 82-95.
----. 1984. K klassifikacii lezginskix jazykov. *Vja* 1984/85,88-94.
----. 1985a. *Voprosy sravnitel'no-istoričeskoj grammatiki lezginskix jazykov*. Moskva: Nauka.
----. 1985b. Struktura prostogo predloženija v lezginskix jazykax. Maxačkala: IzAN *SLiJA* 44,4:309-317.
----. 1994a. Budukh. R. Smeets (ed.) 1994:259-296.
----. 1994b. Rutul. R. Smeets (ed.) 1994:213-258.

ALIPULATOV, M.A. 1974. Kategorija grammatičeskix klassov v jazykax lezginskoj gruppy (po dannym imen). *EIKJ* I:292-297.

BOKAREV, E.A. 1961. *Vvedenie v sravnitel'no-istoričeskoe izučenije kavkazskix jazykov.* Maxačkala: DagUč.

DEETERS, G. 1963. Die kaukasischen Sprachen. *HdO* 7,1.1-79.

DIRR, A.1913. Caxurskij jazyk. Grammatičeskij očerk, teksty, sbornik caxurskix slov s russkim k nemu ukazatelem. *SMOMPK* 43,3.

----. 1928. *Einführung in das Studium der kaukasischen Sprachen.* Leipzig, Asia Major.

DIXON, R.M.W. 1994. *Ergativity.* Cambridge: CUP.

DROSSARD, W. 1982. Nominalklassifikation in den ostkaukasischen Sprachen. H.J. Seiler / Fr.J. Stachowiak (eds.). *Apprehension. Das sprachliche Erfassen von Gegenständen,* Teil II. Tübingen: Narr, 155-178.

ERCKERT, R. v.E. 1895. *Die Sprachen des kaukasischen Stammes. Mit einem Vorwort von Fr. Müller.* Wien.

GENKO, A. 1934. *Caxurskij (caxskij) alfavit.* Baku: Elm.

GIGINEJŠVILI, B.K. 1977. *Sravnitel'naja fonetika dagestanskix jazykov.* Tbilisi: Mecniereba.

HÜSING, G. 1910. Die elamische Sprachforschung. *Memnon* IV.

IBRAGIMOV, G.X. 1968. *Fonetika caxurskogo jazyka.* Maxačkala, Dag.FAN.

----. 1981. Fonomorfologičeskij analiz struktury slova v caxurskom jazyke. B.B. Talibov (otv.red.). *Fonetičeskaja sistema dagestanskix jazykov.* Maxačkala: Dag.FAN, 27-37.

----. 1978a. *Rutul'skij jazyk.* Moskva: Nauka.

----. 1978b. Dialektnaja differencacija caxurskogo jazyka. *IEI* 1976-77.55-56.

----. 1990. *Caxurskij jazyk.* Moskva: Nauka.

K'AXAƷE, O.I. 1984. *Gramat'ik'uli k'lasebi lezgiur enebši - ist'oriul-šedarebiti analizi.* Tbilisi: Mecniereba.

KIBRIK, A.E. 1994. Archi. R. Smeets (ed.). *The Indigenous Languages of the Caucasus,* vol. 4. Delmar, N.Y.: Caravan, 297-365.

---- / S.V. KODZASOV 1990. *Sopostavitel'noe izučenie dagestanskix jazykov. Imja, fonetika.* Moskva: IzdMosUniv.

----. et al. 1977. *Opyt strukturnogo opisanija Arčinskogo jazyka.* Tom II: Taksonomičeskaja grammatika. Moskva: IzdMU.

KLIMOV, G.A. 1977. *Typologija jazykov aktivnogo stroja.* Moskva: Nauka.

KODZASOV, S.V. / I.A. MURAV'EVA 1982. Fonetika tabasaranskogo jazyka. *Tabasaranskie ètjudy.* Moskva: Izd. Moskovskogo Universiteta.

MAGOMETOV, A.A.1970. *Agul'skij jazyk. Issledovanie i teksty.* Tbilisi: Mecniereba.

----. 1980. K èvoljucii klassnoj sistemy v dagestanskix jazykax. *VJa* 1980,1.70-75.

MAXMUDOVA, S. / H. VAN DEN BERG (in preparation). *Rutul.* München: Lincom.

MEL'NIKOV, G.P / A.I. KURBANOV 1964. Logičeskie osnovanija imennoj klassifikacii v caxurskom jazyke. V.V. Ivanov (otv.red.). *Voprosy struktura jazyka.* Moskva: Nauka,157-170.

MIKAILOV, K.Š. 1968. Refleksy avarskix abruptivnyx affrikat v arčinskom jazyke. *Uč.Zap. Dag.fil.AN* 18.201-211.

----. 1972. Ob arčinskix sootvetstvijax avarskim šumnym spirantam. *VDVJa,* 123-134.

MOOR, M. 1985. *Studien zum lesgischen Verb.* Wiesbaden: Harrassowitz.

NICHOLS, J. 1997. The origin of the Chechen and Ingush. Paper presented at the Seventh Conference on the Cultures of the Caucasus. Chicago, May 1997.

NIKOLAEV, S.L. / S.A. STAROSTIN 1994. *A North Caucasian Etymological Dictionary.* Moscow: Asterix.

SCHULZE, W. 1988a. *Protolezgica - Studien zur Rekonstruktion des Lautstandes der süd-ostkaukasischen (lezgischen) Grundsprache.* Bonn (Habil.Schrift Univ. Bonn).

----. 1988b. Noun Classification and Ergative Construction in East Caucasian Languages. *Studia Caucasologica* I,51-274.

----. 1992. Zur Entwicklungsdynamik morphologischer Subsysteme. Die ostkaukasischen Klassenzeichen. C. Paris (éd.). *Caucasologie et mythologie comparée.* Paris: Peeters, 335-362.

----. 1997a. *Person, Klasse, Kongruenz. Fragmente einer Kategorialtypologie des "einfachen Satzes" in den ostkaukasischen Sprachen.* Vol. I: Die Grundlagen. München: Lnicom (to appear).

----. 1997b. *Person, Klasse, Kongruenz. Fragmente einer Kategorialtypologie des "einfachen Satzes" in den ostkaukasischen Sprachen.* Vol. II: Die Person. München: Lincom (to appear).

----. (in preparation). *Person, Klasse, Kongruenz. Fragmente einer Kategorialtypologie des "einfachen Satzes" in den ostkaukasischen Sprachen.* Vol. III: Nominalklassifikation.

SCHULZE-FÜRHOFF, W. [= Schulze] 1992. How can class markers petrify? Towards a functional diachrony of morphological subsystems in the East Caucasian languages. H.I. Aronson (ed.) *The Non-Slavic Languages of the USSR, Linguistic Studies,* New Series. Chicago: CLS, 198-233.

----. 1994a. Zur funktionalen Syntax des "einfachen Satzes" im Xinalug. R. Bielmeier et al. (eds.). *Indogermanica et Caucasica,* Festschrift für K.H. Schmidt. Berlin / New York: De Gruyter, 500-523.

----. 1994b. Udi. R. Smeets (ed.). *The Indigenous Languages of the Caucasus,* vol. 4. Delmar, N.Y.: Caravan, 447-514.

----. 1994c. Tracing Aspect Coding Techniques in the Lezgian Languages. H. Aronson (ed.). *Non-Slavic Languages of the USSR. Papers From The Fourth Conference.* Columbus, Ohio: Slavica, 193-208.

ŠIXSAIDOV, A.R. 1987. Zakarija al-Kazvini o Dagestane. V.G. Gadžiev (ed.). *Istočnikovedenie istorii dosovetskogo Dagestana.* Maxačkala: Dag.fil. AN SSSR.

TALIBOV, B.B. 1960. Mesto xinalugskogo jazyka v sisteme jazykov lezginskoj gruppy. Uč.Zap. *IIJaL* Maxačkala VII,281-304.

----. 1967. Caxurskij jazyk. *JNSSSR* IV,591-607.

----. 1980. *Sravnitel'naja fonetika lezginskix jazykov.* Moskva: Nauka.

TOPURIA, G.V. 1973. K obrazovaniju množestvennogo čisla v lezginskix jazykax. *IKJ* 17. 285-302.

----. 1995. *Morfologija slkonenija v dagestanskix jazykax.* Tbilisi: Mecniereba.

TROMBETTI, A. 1922. *Elementi di Glottologia.* Bologna: Zanichelli.

TRUBETZKOY, N. [Troubetzkoy] 1922. Les consonnes latérales dans les langues caucasiques septentrionales. *BSL* 23.3,184-204.

TRUBETZKOY, N. 1931. Die Konsonantensysteme der ostkaukasischen Sprachen. *Caucasica* 8,1-52.

XAJDAKOV, S.M. 1975. *Sistema glagola v dagestanskix jazykov*. Moskva: Nauka.

----. 1980. *Principy imennoj klassifikacii v dagestanskix jazykax*. Moskva: Nauka.

ǮEIRANIŠVILI, EVG. 1983. *C'aχuri da muχaduri (rutuluri) enebi* Vol. I. Ponet'ik'a. Tbilisi: Izd.TU.

ǮEIRANIŠVILI, EVG. 1984. *C'aχuri da muχaduri (rutuluri) enebi* Vol. II.: Morpologia. Tbilisi: Izd.TU.

Tsakhur is spoken by some 13.000 people who dwell in about 30 villages or settlements at the headwaters of the river Samur (valley of Gorgin Magal) in Southern Daghestan. An important group of Tsakhur speakers can also be found in Northern Azerbajdzhan (along the two tributaries of the Agri-Chay river (Katekh-Chay and Kurmukh-Chay)). "Tsakhur" is the somewhat disputed) name for a dialect continuum, that is named for the village of Tsakhur (in the Samur valley). Together with Rutul, the language forms the western branch of the Samur languages, itself being a subgroup of South East Caucasian (Lezgian). Though Tsakhur is only sporadically written (a new "written language" has recently been reintroduced), it is quite vivid in ordinary life (competing especially with Azeri).
Tsakhur is a "typical" Lezgian language, operating on a system of semantic ergativity and noun classification, based on extensive case marking and a complex verbal paradigms. Though Tsakhur is heavily agglutinating, inflectional features can often be observed. As opposed to some other Lezgian languages, Tsakhur shows a tendency towards personal agreement (restricted, however, to the first person), ergative case marking of personal pronouns, and the development of focus particles.

The booklet informs on the basic structure of Tsakhur (phonology, morphology, and syntax), which is (at least partly) explained on the basis of internal and external reconstruction. The material stems either from written sources or from own field notes. A sample text together with an interlinear interpretation helps to illustrate the linguistic structure of the language.

LINCOM EUROPA
München - Newcastle

ISBN 3 89586 150 2